DOCUMENT RESUME

ED 134 340 PS 009 079

AUTHOR	Taylor, Terry D.; And Others
TITLE	Final Report on Evaluation and Research of the Flexible Learning System: An Early Childhood Teacher Education Program.
INSTITUTION	Far West Lab. for Educational Research and Development, San Francisco, Calif.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Nov 75
NOTE	205p.; For related documents, see ED 129 451-467 and PS 009 078.
EDRS PRICE	MF-$0.83 HC-$11.37 Plus Postage.
DESCRIPTORS	Check Lists; Childrens Literature; Classroom Environment; *Early Childhood Education; *Evaluation Methods; Performance Based Education; *Performance Based Teacher Education; Preschool Education; *Program Evaluation; Psychomotor Skills; Teacher Education Curriculum; *Teacher Evaluation; *Teaching Skills; Testing
IDENTIFIERS	*Flexible Learning System

ABSTRACT

This is the final report on evaluation and research of the Flexible Learning System (FLS). The Flexible Learning System (FLS) began as an outgrowth of the Far West Laboratory's several years of investment in research and development of a Responsive Model of Early Childhood Education for young children. The model emphasizes the development of cognitive abilities and positive self-concept in children as complementary goals. The principal objective of FLS has been to provide a flexible system of teacher education for adults in early childhood education. This report presents a research and evaluation study of the FLS training materials' effectiveness. Included are the following sections: (1) introduction to FLS, containing an historical overview, description of the FLS, FLS training process and selection of FLS content areas; (2) evaluation of the FLS, containing preliminary field test and results, and main field test and results; (3) operational field test, including design, subjects and procedures; (4) operational field test measurement, covering instrumentation paper-pencil method, work sample method and stimulus materials; (5) definition of teaching competencies, including setting up the physical arrangement of the classroom; providing enriching literature experiences for children, and developing children's sense perception; (6) operational field test results, including effects of physical arrangement, literature experiences, and sense perception treatment, as well as pre- and post-training results. The last section is a discussion of findings. Appendices include instruments used, reliability figures and supplementary data. (HS)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original.

ED134340

12176

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Final Report on Evaluation and Research of the Flexible Learning System: An Early Childhood Teacher Education Program

Terry D. Taylor
Judith J. Hubner
Stanley H.L. Chow
Steven R. Stripp
Masako N. Tanaka

November 1975

PS 009079

FAR WEST LABORATORY FOR EDUCATIONAL RESEARCH AND DEVELOPMENT

This learning unit is a product of the Early Childhood Education Program of the Far West Laboratory for Educational Research and Development, San Francisco, a public, nonprofit organization supported in part by funds from the National Institute of Education, Department of Health, Education and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement should be inferred.

Far West Laboratory for Educational Research and Development

1855 Folsom Street, San Francisco
California 94103 - (415) 565-3000

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

Far West Laboratory for Educational Res. & Dev.

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

Copyright © 1975 Far West Laboratory for Educational Research and Development. All rights reserved. Copyright for these materials is claimed only during the period of development, test, and evaluation, unless authorization is granted by the National Institute of Education to claim copyright also on the final materials. For information on the status of the copyright claim, contact either the copyright proprietor or the National Institute of Education.

ED134340

Final Report on Evaluation and Research of the Flexible Learning System: an Early Childhood Teacher Education Program

Terry D. Taylor

Judith J. Hubner

Stanley H. L. Chow

Steven P. Stripp

Masako N. Tanaka

November, 1975

Far West Laboratory for Educational Research and Development

San Francisco, California

PS 009079

Table of Contents

	Page
Acknowledgements	i
List of Figures	ii
List of Tables	iii
Introduction to the Flexible Learning System	1
Historical Overview	1
Description of the Flexible Learning System	3
The Flexible Learning System Training Process	4
Selection of Flexible Learning System Content Areas	5
Evaluation of the Flexible Learning System	8
Preliminary Field Test	11
Preliminary Field Test Results	12
Main Field Test	12
Main Field Test Results	17
Operational Field Test	21
Design	22
Subjects	23
Procedures	25
Operational Field Test Measurement	27
Instrumentation	27
Paper-Pencil Method	27
Work Sample Method	28
Stimulus Materials	28

Defining Teaching Competencies 31

Physical Arrangement of the Classroom 34

Enriching Literature Experiences of Children 36

Developing Children's Sense Perception 36

Operational Field Test Results 39

Treatment Pretraining Results 41

Physical Arrangement Treatment Effects 43

Literature Experience Treatment Effects 43

Sense Perception Treatment Effects 47

Educational Pre- and Posttraining Results 52

Discussion of Findings 62

Effects of Training 62

Summary of Training Effectiveness 65

Effects of Educational Level and Methods of Measurement 67

Summary of High and Low Education Teacher's Performance 69

Recommendations for the Flexible Learning System 72

List of Appendices

A. Paper-Pencil and Work Sample Instruments, Definitions, and Procedures

B. Reliability for Work Sample and Paper-Pencil Measures

C. Supplementary Data: Methods and Results

Acknowledgments

A large number of people have invested energy, time, and patience in this report and the research which it presents. The teachers and their supervisors who participated in the study suffered disruption and confusion with little complaint. Bill McGuiness, director of the San Diego Head Start Program, handled coordination problems and emergencies that would have overtaxed most persons. The FLS research assistants Seena D. Alenick, Carole Brown, Louise Caraco, Esther Ehrlich, Carol Rice, and Renee Zarlow gave meaning to the word competency by their careful and sensitive formulation of competency definitions.

Special thanks is due Carol Rice for her contributions to the entire FLS evaluation effort. Cecilia Culang also deserves special praise for her willingness to work extra hours to type, retype, and type again a report that she must now have memorized.

List of Figures

Figure		Page
1.	Assessment Levels in Teacher Education.....	9
2.	A conceptual hierarchy of the Flexible Learning System.....	32
3.	Competency hierarchy for the module Physical Arrangement of the Classroom.....	35
4.	Competency hierarchy for the module Enriching Literature Experiences of Children.....	37
5.	Competency hierarchy for the module Developing Children's Sense Perception.....	38

List of Tables

Table		Page
1a.	Results of PFT Learner Feedback on Open-ended Questions	13
1b.	Results of PFT Learner Feedback Questionnaires by Percentages of Positive Ratings	14
1c.	Results of PFT Learner Feedback on a Rating Scale Questionnaire	15
2.	Main Field Test Pre- and Posttest Mean Scores for Seven FLS Modules	16
3.	Sample Size, Mean, and Mean Difference Scores Between Pretest and Posttest Performance of Highly Educated Trainees and Highly Experienced Trainees	19
4.	Demography of Study Sample by Total and Hi-Low Education Groups	24
5.	Research Design	26
6.	Distribution of Possible Points Within the Paper-Pencil Measure Subscales	29
7.	Pretraining Performance Means, Standard Deviations and F Values for the OFT Training Groups	42
8.	Pre- to Posttraining Performance Means, Standard Deviations, and F Values for Physical Arrangement Treatment Group	44
9.	Posttraining Performance Means, Standard Deviations, and t Values for the Physical Arrangement Treatment and Control Groups	45
10.	Posttraining Performance Means, Standard Deviations, and t Values for the Physical Arrangement Treatment Group vs. Each Other Group	46

11. Pre- to Posttraining Performance Means, Standard Deviations, and F Values for Literature Awareness Treatment Group..... 48

12. Posttraining Performance Means, Standard Deviations, and t Values for Literature Awareness Treatment and Control Groups..... 49

13. Posttraining Performance Means, Standard Deviations, and t Values for the Literature Experience Treatment Group vs. Each Other Group..... 50

14. Pre- to Posttraining Performance Means, Standard Deviations, and F Values for Sense Perception Treatment Group..... 51

15. Posttraining Performance Means, Standard Deviations, and t Values for the Sensory Perception Treatment and Control Groups..... 53

16. Posttraining Performance Means, Standard Deviations and t Values for the Sense Perception Treatment Group vs. Each Other Group..... 54

17. Pretraining Performance Means, Standard Deviations, and F Values Between Two OFT Training Groups Differing in Educational Background..... 55

18. Posttraining Performance Means, Standard Deviations, and t Values for High Education vs. Low Education Groups that Received Training..... 57

19. Posttraining Performance Means, Standard Deviations, and t Values for Physical Arrangement High vs. Low Education Teachers..... 58

20. Posttraining Performance Means, Standard Deviations, and t Values for Literature Experience High vs. Low Education Teachers..... 60

21. Posttraining Performance Means, Standard Deviations, and t Values for Sense Perception High vs. Low Education Teachers..... 61

THE FLEXIBLE LEARNING SYSTEM

The Flexible Learning System (FLS) began as an outgrowth of the Laboratory's several years of investment in research and development of a Responsive Model of Early Childhood Education for young children. The System's principal objective has been to provide a flexible system of teacher education for adults in Early Childhood Education. This report is a record of that growth and presents a research and evaluation study of the FLS training materials' effectiveness.

Historically the FLS traces its origins at the Far West Laboratory to the start of the Responsive Education, Head Start and Follow Through programs. As early as 1966 when the Laboratory was established and supported by ESEA Title IV, the Responsive Model of early education for preschool children had been developed and tried out on a small scale at the New Nursery School in Greely, Colorado. In 1967, its originator, Dr. Glen Nimnicht, joined the Laboratory and began operating a national Head Start program using the Responsive Model. In 1968, the national focus on early childhood education expanded to include the children from Kindergarten through the third grade. Follow Through was launched as an attempt to follow-up whatever benefits Head Start had accrued for young children and to expand early stimulation and enrichment programs beyond the preschool ages. In the same year, the Responsive Education program was established as a sponsor for a score of Follow Through Sites. At present, the Responsive Education staff is engaged in installing, maintaining and evaluating the program at 14 sites across the country.

The assumptions of the Responsive Model are that children learn at

different rates and in different ways, and above all, they learn best when they are interested and motivated to learn. Therefore, the school environment must make provisions for these differences and respond to the needs of individual children. The model emphasizes the development of cognitive abilities and positive self-concept in children as complimentary goals for a successful early childhood program. A responsive classroom is one in which children are free to explore the environment. Activities are self-rewarding. Children work at their own pace, receive immediate feedback, and are encouraged to make their own discoveries.

One of the major concerns of the Head Start and Follow Through Programs was and continues to be early education staff development. The need for additional trained early childhood staff was evident through the early seventies - witness the Child Development Associate (CDA) program which was launched to train and certify paraprofessionals in child care services. It was felt that the experience in training of the Responsive Model staff during the past several years presented a valuable source of expertise from which to develop an exportable training system. In 1972, Dr. Nimnicht received funds from the National Institute of Education to develop and evaluate a prototype system to prepare paraprofessionals for careers in early childhood education. The Flexible Learning System was initially conceived of as a competency-based training system consisting of a number of training units designed for early childhood staff members.

Description of the FLS

Currently, the Flexible Learning System consist of an initial set of 15 learning modules for adults in early childhood education. It is intended that these modules provide learner-centered, functional-context, competency-based experiences.

The FLS is learner-centered at three levels: at entry into the system, during the process of training, and at exit from the system. Each potential trainee who enters the system has the opportunity to select modules of personal interest. During training, teachers have the opportunity to try out various roles of persons involved in the early childhood educational environment. For example, they are provided experiences in which they play the child role, as well as experiences in which they practice teaching skills with other adults who provide feedback. When learners exit from training, they should be prepared to extend their learnings beyond the module, to "personalize" their learning to fit their individual setting, children's needs, personal, and programmatic teaching goals.

The FLS provides a functional-context experience in that it is designed for learning while on-the-job. Each module provides opportunities for daily application and practice with young children aged four to nine. The system is competency-based in that it is based on the doing of teaching acts, not just reading or hearing about them. Each of the training modules has explicitly stated objectives. The competent teacher is one who masters these objectives in order to work effectively with children. This requires behavioral skill, not just conceptual understanding. At present, outcome competencies have been identified for each module in general terms; for example, 'Know how to

change the number and location of classroom areas to suit your own needs and those of the children; or, to enable the teacher to develop his or her own curriculum for sense perception using checklists of behavioral (or other types of) objectives, wordlists, sensory activities, diagnosis, observation, and individualized follow-up (prescription); and, learning to apply the guidelines to select books for the purposes of this module--using that one way as a model--.'

The FLS Training Process

Although each of the 15 modules of the FLS was developed in a unique style, reflecting both the nature of the subject matter and the personal style of the developer, there are six procedural steps common to the training process used in most of the modules.

First, each begins with an introduction which includes performance objectives, learning methods, and learning materials and estimates as to the length of time needed to complete the module. The second step involves a conceptual presentation of the subject matter of the module. This may involve studying the module content and participating in discussions led by the trainer. The third step is aimed at conceptual integration. Activities are presented which require personalized application of the basic concepts and ideas in the module. Once the groundwork of conceptual understanding has been laid, the fourth step requires behavioral production of particular teaching skills. This may involve role-playing sessions with other group members, or trying out the new concepts in devising lesson plans. The fifth step is classroom application. At this point, learners are provided with activity guides that enable them to apply the skills they have learned while working with children in their own classrooms. The

sixth step is intended as the final outcome of each of the modules. It involves the learner's "personalization" of the training experiences for application to his/her own situation. This requires a "thinking through" of personal teaching goals and a revision and extension of FLS training to facilitate achievement of them.

Criteria are described for self-evaluation throughout the learning process and for determining the levels of competence obtained after studying the module. Verbal recall is the first level of competence and may be demonstrated either by written production or observational identification of a competency. The second level of competence requires the production of a learned behavior. The third level can be demonstrated by the generalization of the learned behavior, for example, creating a new learning episode. The fourth level of competence is demonstrated by going beyond what was learned in the module by taking ownership of the skills developed in the module and developing a personally relevant form of expressing them. Finally, the fifth level involves demonstration of ability to teach to others the skills in a module, and is usually required of a trainer.

Selection of FLS Content Areas

The FLS modules as they now exist constitute a prototype training system for early childhood staff. These 15 modules represent an initial array of skills which may be considered a starter set of teacher competencies upon which further staff growth and development may be built.

In selecting and determining the content for this beginning set of modules, the staff was influenced by two sources of competency taxonomies. The primary source was the original competency matrix developed by Nimnicht

(Nimnicht & Johnson, 1973) which consisted of 44 training units clustered around three grade spans (infant, preschool, K-3rd) and four staff roles (assistant, teacher, Program advisor, and Program director). A further influence was the competency taxonomy developed for the Child Development Associate Program (OCD, 1973). The CDA taxonomy included 42 competencies organized around six competency clusters.[1] An examination of these two taxonomies reveals substantial overlap in certain competency areas. The selection of modules to develop within the confines of a prototype FLS system was based to some extent on the degree of overlap and agreement between these areas; thus, the FLS modules represent to some degree a subsample of each taxonomy. During the course of the FLS program, these modules were sorted into three clusters: (a) Understanding the Child, (b) Organizing the Learning Environment and, (c) Facilitating Children's Learning.

---

[1]These six clusters refer to 1. Setting up and maintaining a safe and healthy learning environment, 2. advancing physical and intellectual competence, 3. building positive self-concept and individual strength, 4. organizing and sustaining positive functioning of children and adults in a group, 5. bringing about optimal coordination of home and center childrearing practices, 6. carrying out supplementary responsibilities.

CLUSTER I: Understanding the Child is concerned with skills related to the teacher's sensitivity and proficiency in understanding the child's social and psychological needs and in determining his/her level of intellectual development. It attempts to include strategies that can enable teachers and aides to determine which stage of social or cognitive development the child is at and what knowledge and skills he/she has brought to the learning environment of the classroom. Modules developed within this cluster include:

1. Assessing Children's Cognitive Progress
2. Communicating and Working with Parents
3. Developing Children's Self-Concept
4. Observing and Extending Spontaneous Play

CLUSTER II: Organizing the Learning Environment includes many of the management planning and coordination skills needed in operating an early childhood program. Its concern is the structuring and provisioning of the learning environment to accommodate needs of children. Modules developed within this cluster include:

1. Arranging the Physical Environment
2. Managing the Preschool Classroom
3. Overview of Early Childhood
4. Selecting Children's Books with a Black Perspective
5. Content Analysis of Materials about Chicanitos

CLUSTER III: Facilitating Children's learning includes specific skills that can promote the development of the child's intellectual capacity. In addition, it addresses early childhood curriculum concerns. This cluster consists of the modules:

1. Teaching Language in an Integrated Way

2. Enriching Children's Literature Experiences

3. Problem-Solving with Children

4. Developing Sense Perceptions in Children

5. Using Toys and Games with Children

6. Developing Children's Concept Formation Skills

Evaluation of the Flexible Learning System

A three phase field test cycle was established for the purpose of evaluating the modules of the Flexible Learning System during and after their development. The first phase of evaluation was the Preliminary Field Test (PFT). The module developer delivered training and trainee feedback as to the feasibility, strengths, and weaknesses of each competency module was collected to allow revision of it. The Main Field Test (MFT) comprised the second phase of competency module evaluation. Generally speaking, the test conditions for this phase were designed to allow a more independent field implementation of the module than the PFT permitted and to assess trainee performance as an outcome of training. On site personnel delivered training and pre- and post- paper-pencil examinations were administered to trainees at the first and last training sessions as an assessment of competencies delivered by each module. The Operational Field Test (OFT) was the third phase of field testing. Data was collected which focused on the question of whether modules produce desirable changes in teacher behavior in the natural setting. This focus is in line with the sequence of levels of assessment proposed by Medley, Soar, and Soar (1975). They proposed a simple paradigm for competency evaluation consisting of four assessment levels (see Figure 1).

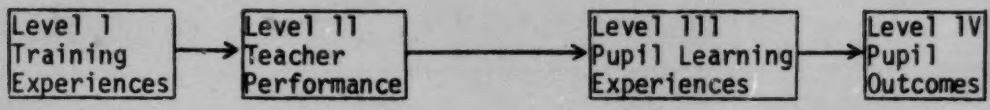


Figure 1. Assessment Levels in Teacher Education
(Medley, Soar, and Soar, 1975, p.2)

Level I refers to assessments of the training experiences the teacher has had: What courses has he taken? What modules has he attempted? Which has he bypassed on the basis of having demonstrated mastery of its objectives beforehand?

Level II refers to assessments of the teacher's behavior while he is attempting to fulfill the role of a teacher. What kinds of questions does he ask in interaction with pupils? How does he organize his class for instruction? How does he determine instructional objectives?

Level III refers to assessments of the behaviors of pupils under the guidance of the teacher being assessed -- assessments of the experiences they have which we all know must form the basis for any learning that takes place. What kinds of tasks do the pupils perform, in or out of class? How much time do they spend in active participation in class discussions? How often does each child receive reinforcement and for what?

Level IV refers to assessments of the outcomes of instruction -- of those changes in behavior that it is the purpose of education to bring about. How well does the pupil read? What are his attitudes toward independent learning in adult life? What kind of a citizen does he become?

Each of these four assessment levels is represented in the diagram by a rectangle and the rectangles are joined by arrows representing lines of influence or of cause-and-effect. Thus, outcomes (Level IV) are seen as influenced by, and in part are the result of learning experiences

the pupil has while in school (Level III); these experiences in turn are seen as at least partly determined by what the teacher does (Level II); and, finally, the way the teacher behaves---"teaches"---is affected by the experiences he has had during training (Level I). All of these things are strongly influenced by other factors not shown in the figure, e.g., community, school, pupil, and teacher characteristics.

The whole enterprise of teacher education is, of course, based on the assumption that, despite any extraneous factors, the influences that are assessed at each level are potent enough to have an appreciable effect not only on the level immediately following but on all subsequent levels. The concept of teacher effectiveness, in particular, is based on the notion that pupil learning outcomes (Level IV) are affected by teacher behavior (Level II). And the justification for the very existence of teacher education is the presumption that what happens to a teacher in training (Level I) can somehow increase his effectiveness, that is, affect pupil learning outcomes (Level IV).

Ideally, evaluation of FLS training effectiveness would encompass all four levels of assessment included in the paradigm. Unless, however, the resources exist for simultaneous assessment at all four levels, the paradigm suggests that the most productive approach to evaluation is sequential assessment at each level. In other words, assessment of training experiences (Level I) is the logical place to begin evaluation of a teacher training process. The next step is to evaluate the teacher performance (Level II) of teaching skills. Then assessment of the effects of teacher performance on pupil learning experiences (Level III) should be accomplished, and finally, pupil outcomes (Level IV) should be measured.

The PFT and MFT evaluations of the FLS focused primarily on Level I

of the paradigm. Extensive data was collected on trainees' background characteristics, training, and mastery of module objectives. These data provided the basis for the OFT focus on Teacher Performance (Level II) assessment as a means of evaluating teacher training effectiveness. This takes competency measurement beyond "what do teachers know" to "how do they use what they know?" (Ward, Morine, and Berliner, 1975). Future research should address the question of how do pupil experiences and outcomes change as a result of teachers using what they know.

Preliminary Field Test

The design of the Preliminary Field Test (PFT) involved the module developer serving as trainer for a small group of early childhood teachers. The form of the module was typically a detailed outline or written summary. Training time varied from 6 to 12 weeks according to the amount necessary to complete each module. The major outcome of the PFT was its service as a formative feedback tool for improving module development through recycling. User feedback data on each module was channeled to the developer for revision of the module. The instruments used for this purpose included a biographical data card for obtaining background information about users, different forms of a structured feedback instrument (the different forms reflected changes toward increasing objectivity and clarity in FLS evaluation) to provide a quantitiative rating of module quality and feasibility for the field, and an open-ended response questionnaire regarding suggestions, revisions, etc., for the module unit. Module quality was defined in a number of ways, e.g., likelihood of activities to accomplish module objectives, clarity of directions, understandability, and comparability with other forms of training. Tables 1a through 1c

present a summary of PFT findings according to the form of feedback instrumentation used. Table 1a shows findings for modules rated using positive or negative ratings; Table 1b presents findings for modules using an instrument with a dichotomous rating scale (yes or no), and Table 1c presents findings for modules rated with a five-point rating scale.

PFT Results

The majority of the modules received positive ratings. Most trainees (90%) rated the modules as interesting, clear, easy to understand and well-organized. Most participants (79%) indicated that the materials were about the same or slightly superior as compared to other training materials, and would recommend the modules to others. These data strongly suggest that the preliminary forms of the modules were feasible in terms of user satisfaction with their scope and direction.

Main Field Test

The design of the Main Field Test (MFT), involved the delivery of each module to a fairly large sample of teachers ($\bar{X}_n$ =31.3). Training was conducted by an on-site trainer/coordinator trained by the module developer in an intensive two-day workshop covering module content and delivery skills. The module was presented in as complete a form as possible. Pre- and posttraining paper and pencil tests were administered at the first and last training sessions to assess written performance on competencies delivered by each module. In addition, background information was collected using a biographical data card, and, when time permitted, feasibility data was collected using the PFT instruments. Most modules were field tested by trainees enrolled in courses offered for academic credit through various academic institutions. A summary of the results of pre- and posttest performance for seven of the modules are presented in Table 2.

Table 1a. Results of PFT Learner Feedback on Open-ended Questions(1972 Form)

Competency Module	Sample Size	Personal Satisfaction with Module						Dissemination Questions		
		Positive Interest	Objectives Clear	Reading Level Easy	Instructions Clear	Function Helpful	Organization Good	Compared Favorably With Other Training	Would Recommend To A Friend	Average Total Rating + = Positive - = Negative
1. Analyzing Materials for Use by Chicanitos	11	+ *	+	+	+	+	+	+	+	+
2. Helping Children Develop a Healthy Self-Concept	28	+	+	+	-	n.d.**	+	n.d.	n.d.	+
3. Physical Arrangement of the Classroom	10	+	+	+	+	+	+	+	+	+
4. Understanding Children's Thinking	11	+	n.d.	+	+	n.d.	n.d.	n.d.	n.d.	+

* Positive ratings by trainees on open-ended items

** n.d. = no data available

Table 1b. Results of PFT Learner Feedback Questionnaires by Percentages of Positive Ratings (1973 Form)

Competency Module	Personal Satisfaction with Module							Dissemination Questions		
	Sample Size	Positive Interest	Objectives Clear	Reading Level Easy	Instructions Clear	Function Helpful	Organization Good	Compared Favorably With Other Training	Would Recommend To A Friend	Average Total Rating
1. An Introduction to Early Childhood Education and the Responsive Program	16	92%	92%	80%	100%	100%	86%	69%	n.d.*	88%
2. Using Toys and Games to Help Children Learn Specific Skills and Concepts	15	93%	92%	100%	93%	93%	55%	84%	n.d.	87%
3. Developing Children's Senses and Perceptions	7	100%	100%	100%	100%	100%	100%	66%	66%	92%
4. Developing Children's Class Concept-Formation Ability	17	76%	55%	62%	40%	51%	49%	50%	n.d.	55%
5. Developing Children's Problem Solving Ability	36	100%	94%	93%	95%	100%	100%	n.d.	n.d.	97%
6. Analyzing Materials for and about Blacks	6	100%	n.d.	100%	50%	100%	75%	n.d.	n.d.	85%

*n.d. = no data available

Table 1c. Results of PFT Learner Feedback on a Rating Scale Questionnaire (1974 Form)

		Personal Satisfaction with Module						Dissemination Questions		
	Sample Size	Positive Interest	Objectives Clear	Reading Level Easy	Instructions Clear	Function Helpful	Organization Good	Compared Favorably With Other Training	Would Recommend To A Friend	Average Total Rating
1. Teaching Language Skills in an Integrated Way	12	1.25*	1.25	1.37	1.71	1.14	1.57	2.14	1.62	1.51
2. Literature Experience for Young Children	16	2.12	2.19	1.62	1.69	1.94	1.73	2.33	2.12	2.0
3. Classroom Management	17	1.7	1.43	1.47	1.71	1.80	1.93	2.53	1.47	1.76
4. Utilizing the Spontaneous Play of Children	14	1.5	1.2	1.4	1.5	1.4	1.3	n.d.**	1.2	1.4

*Rating scale of 1 to 5 with 1 representing the highest positive rating

*n.d. = no data available.

Table 2. Main Field Test Pre- and Posttest Mean Scores for Seven FLS Modules.

	Sample Size	Average Pretest Score	Average Posttest Score	Mean Difference	Total Score Possible
1. Analyzing Materials for use by Chicanitos	19	52.6	57.2	4.6	89
2. Teaching Language Skills in an Integrated Way	38	24.2	31.8	5.6**	39
3. Helping Children Develop a Healthy Self-Concept	17	69.8	68.0	-1.8	98
4. Developing Children's Problem-Solving Ability	46	19.3	28.4	9.2**	47
5. Physical Arrangement of the Classroom	32	10.4	11.5	1.1	14
6. Literature Experience for Young Children	24	16.7	19.8	3.1**	33
7. Classroom Management	51	55.4	66.5	11.1**	99

*$p < .05$

**$p < .01$

MFT Results

Findings from the MFT of competency modules indicated that there were significant gains between pre- and posttest total scores for four of the seven modules that were field tested; Classroom Management, Developing Children's Problem Solving Ability, Teaching Language Skills in an Integrated Way and Literature Experiences for Young Children. Two modules, Analyzing Materials for Use by Chicanitos and Physical Arrangement of the Classroom, showed significant gains for selected subtest scores; however, their total score gains were not significant.

There were no significant gains in trainee performance following training in the Helping Children Develop a Healthy Self-Concept module. The lack of significant gain in performance appears to be a function of the assessment instrument rather than the training module. The average pretest score was 69.8 out of a possible 98 dichotomous items. If trainees had merely been guessing on the pretest, the average score, based on chance, would have been 46.8. The relatively high average pretest score thus suggests a "ceiling effect." Since trainees were clearly performing above the chance level on the pretest, the test may have been assessing knowledge that the trainees possessed before training, allowing little room for improvement as a result of training. The paper-pencil examination is now in the process of revision and will assess a larger array of competencies when it is completed.

Correlations between selected trainee characteristics and pre- and posttest mean scores were generated from the MFT data for each module. Educational background of trainees appeared to be particularly important in terms of its relationship to trainee paper-pencil exam performance in the Main Field Test. The data suggested that less educated trainees tended to have more years of experience in working with young children but consistently scored lower on the pre- and posttests than their younger colleagues with more formal education.

To examine this issue in more detail, MFT data for three modules underwent further analysis. Table 3 presents the mean difference scores between pre- and posttest performance for these modules. Note that the high experience (low education) trainees scored lower than the high education trainees, and that their pre- to posttest gain scores were significant ($p < .05$) for two of the three modules.

There are at least two possible explanations for this. First, the MFT assessment method may have accurately measured differences in competence of teachers with different levels of education. Teachers with more formal education may be more competent than less educated teachers. The alternative explanation is that the MFT assessment method confounded teaching competence and paper-pencil test-taking competence. That is, less educated trainees may not score as high on paper-pencil exams as trainees with more education because of verbal ability, reading and writing skills, or "test-wiseness". However, since older trainees tend to have more years of work experience it is possible that they could better demonstrate competencies in ways other than just paper-pencil test performance. Given their teaching experience, they may well perform better in applied classroom situations than on written tests. Thus,

Table 3. Sample Size, Mean, and Mean Difference Scores Between Pretest and Posttest Performance of Highly Educated Trainees and Highly Experienced Trainees

	High Education				High Experience			
	n	Pretest Mean	Posttest Mean	Mean Difference	n	Pretest Mean	Posttest Mean	Mean Difference
1. Classroom Management	18	61.38	71.61	10.23**	4	41.75	55.50	13.75*
2. Teaching Language Skills in an Integrated Way	12	27.41	33.91	6.50**	6	24.66	31.50	6.84*
3. Literature Experience for Young Children	10	17.90	21.80	3.90**	2	13.00	17.00	4.00

* $p < .05$

** $p < .01$

early childhood teacher evaluation should pursue alternative methods of assessing teacher performance.

The Flexible Learning System does not stand alone in its concern with appropriate methods of assessing the effects of teacher training. A growing body of professional educators have advocated that research on teacher performance should involve the assessment of classroom behavior to evaluate teacher effectiveness (e.g., Dunkin and Biddle, 1974). It is argued that a teacher's performance in a reality-based situation with live students is a more valid indicator of teacher training effectiveness than the standard paper-pencil tests which traditionally have been a part of the typical teacher education curriculum. Thus, to ensure that teachers are fairly evaluated, it may be necessary to use additional methods of assessment such as observation of classroom behavior or supervisor judgments.

To summarize, research on the effects of teacher training has suggested the usefulness of observational performance assessment as an evaluation method. However, traditional teacher training evaluation has relied only on paper-pencil assessment; while research studies have relied almost exclusively on performance assessment. For example, Borg (1972) reported that after 39 months, teachers who completed a mini-course designed to help develop nine skills and extinguish three undesirable behaviors related to the discussion method of teaching, were observed in a teaching situation and performed at a significantly higher level than they did prior to the training.

There are two lines of argument that suggest reliance on either paper-pencil or measures of classroom behavior may be an inappropriate strategy in studies of early childhood teacher performance. Teachers

are characterized by a wide variety of experiences, and educational levels. As suggested by the PFT and MFT data, differences in background characteristics may tend to confound teaching competence with paper-pencil test-taking competence. Therefore, it may be desirable to use multiple methods of measurement to discriminate between performance of teachers who diverge in characteristics such as amount of education or work experience.

The second argument against relying on a single method of measurement is provided by Campbell and Fiske (1959). They maintain that alternative methods should be applied to measure the same construct, or trait, as an indication of convergent validity of measurement. This suggests the importance of using both paper-pencil and performance assessment. In addition, Messick (1975) has argued that educational assessment should also focus on construct validation; that is, identification of what is being measured, and empirical validation of that construct. Use of multiple methods to measure teacher performance can establish construct validation by generating estimates of validity through triangulation.

Operational Field Test

The findings from the Preliminary and Main Field Tests were instrumental in focusing the design of the Operational Field Test (OFT). While the PFT and MFT provided substantial indication that the FLS training program could deliver knowledge of competencies necessary for teaching in the early childhood setting, they did not answer the question of whether FLS training delivered skill at classroom application of these competencies and whether there were limitations to using paper-pencil instruments as the primary method of competency assessment of teachers

with diverse background characteristics. As a result, the OFT was designed to examine whether:

1. early childhood teachers and assistants trained on a specific FLS module score higher on module-related paper- and pencil and classroom performance measures than those who receive training on other modules or no training at all, and,
2. teachers with a high formal education background (BA degree) perform differentially on paper-pencil and classroom performance measures as compared to teachers with low formal education background (high school diploma).

OFT Design

To produce a research design for addressing these questions, a panel of research and evaluation experts was convened in November, 1974. In addition to the FLS evaluation staff, participants in that panel included: Dr. David Berliner, Far West Laboratory; Dr. Leonard Cahen, Far West Laboratory; Dr. Steven Klein, UCLA; Dr. Virginia Shipman, ETS; and Dr. Gary Borisch, University of Texas. This panel was charged with recommending an optimal research design which would answer the research and evaluation questions listed above. Several alternative designs were proposed and discussions centered around the pros and cons of each, given the available resources.

A central issue faced by the research panel was to decide how many, if not all, of the fifteen FLS modules should be involved in the research effort. The decision reached was that three modules would be investigated. This number was small enough to allow an intensive

examination of their effects and to accomodate the limited resources available yet large enough to be fairly representative of the entire array of FLS modules.

The three modules selected were (a) Physical Arrangement of the Classroom, (b) Literature Experiences for Young Children, and (c) Developing Children's Sense Perception. They were chosen because of their independence from one another in terms of content and because they sample important domains in early childhood education. This increased the representatives of the modules in regard to the total FLS as well as provided an indication of the usefulness of the training process and content embedded in separate modules.

Subjects

A subsample of 32 teachers were drawn from a sample of 110 preschool teachers and assistants employed in Head Start or Children Center Programs in San Diego, California. All participants had the autonomy as well as the opportunity to implement changes in classroom arrangement if they so desired. Due to the attrition of one teacher in the sample, the final sample of the OFT consisted of 31 teachers and assistants.

The teachers were selected on the basis of academic degree attained: 15 participants had high school diplomas and 16 participants had B.A. degrees. Table 4 presents a breakdown of background characteristics of the teachers in the study before and after they were assigned to a high (B.A) or low (high school diploma) education group. Almost all study participants had some formal education beyond high school. One third of the participants (32%) had 14 years of schooling, one third (32%) had 16 years of schooling and one third (32%) had 17 or more years of schooling.

Table 4. Demography of Study Sample by Total and Hi-Low Education Groups.

	Total Group (M=31)	Hi - Ed. (n=16)	Lo - Ed. (n=15)
Position: teacher	27 (87%)	16 (100%)	11 (73%)
teaching assistant	4 (13%)	0 (0%)	4 (27%)
Age: under 30	18 (60%)	13 (81%)	5 (36%)
30-45	8 (27%)	3 (19%)	5 (36%)
over 45	4 (13%)	0 (0%)	4 (28%)
Sex: male	4 (13%)	1 (6%)	3 (2%)
female	27 (87%)	15 (94%)	12 (98%)
Parent: yes	12 (39%)	3 (19%)	9 (60%)
no	19 (61%)	13 (81%)	6 (40%)
Teaching Credential: yes	15 (48%)	13 (81%)	2 (13%)
no	16 (52%)	3 (19%)	13 (87%)
* Use: teach	29	15	14
train	5	3	2
other	1	0	1
* Children: pre-K	30	15	15
Work K-3	5	4	1
other	1	1	0
Years of Education:			
6 yrs or less	0 (0%)	0 (0%)	0 (0%)
12 yrs or less	1 (3%)	0 (0%)	1 (7%)
14 yrs or less	10 (32%)	0 (0%)	10 (67%)
16 yrs or less	10 (32%)	6 (38%)	4 (26%)
17 yrs or more	10 (32%)	10 (61%)	0 (0%)
Years of work experience:			
1 yr or less	6 (20%)	2 (12%)	4 (27%)
3 yrs or less	13 (43%)	10 (62%)	3 (20%)
8 yrs or less	9 (30%)	3 (19%)	6 (40%)
15 yrs or less	0 (0%)	0 (0%)	0 (0%)
15 yrs or more	2 (7%)	1 (7%)	1 (13%)

* several participants checked more than one response.

Twenty-seven participants were teachers and four were teaching assistants. Eighteen of the participants were under 30 years old; 8 were between 30-45 years old; 4 were over 45 years old. The sample consists of mostly female teachers (27) who were working with pre-kindergarten students. About half of the sample (15) had teaching credentials; less than half of the sample (12) were parents. Almost all participants had one to eight years of teaching. Six teachers had one year experience; thirteen had 3 years or less experience; and nine had 8 years or less experience.

The high-low education subsamples differed on several demographic variables. The high education teachers, as a whole, tended to be younger than the low education teachers (81% under 30) and were non-parents possessing teaching credentials. They also tended to have slightly less working experience with children. For example, 74% of the high education group had three years of working experience or less while 53% of the low education group had eight years experience or more. On other variables, such as teaching position, sex, and children they worked with, the groups were fairly equivalent.

Procedures

As mentioned earlier, the 32 teachers were blocked on educational level (16 teachers with a BA degree or above, 16 with a high school diploma or less). Teachers within the two groups were then randomly assigned to one of the three treatments or a control condition, resulting in four high education teachers per condition.

In order to obtain a fairly intense look at the effect of the three modules on trainee performance, a repeated-measure factorial design combining a pre-post control group with a time-series data collection technique (Glass, Willson, and Gottman, 1972) was adopted. Graphically the design was as follows:

Table 5. Research Design (n=32)

		2 weeks		6 weeks	2 weeks	
Treatment Group 1	(n=8)	O_1	O_2	T_1	O_3	O_4
Treatment Group 2	(n=8)	O_1	O_2	T_2	O_3	O_4
Treatment Group 3	(n=8)	O_1	O_2	T_3	O_3	O_4
Control Group	(n=8)	O_1	O_2		O_3	O_4

O_n= Observations

T_n= Treatments

Each of the three treatment groups received six weeks of training from an instructor trained by Far West Laboratory staff. Training occurred weekly in 1-1/2 to 2 hour sessions. Three weeks before training began, all teachers received a set of instructions explaining that they would be observed twice before training and twice arter, and indicating how they were to prepare for these observations. At the same time, each teacher received a set of instructional materials for preparing a sensory activity and a reading activity to be taught on the observation occasions. Data on teacher performance during these activities as well as the physical arrangement of the classroom were collected on all of the teachers. Thus, the Physical Arrangement treatment group was observed on competencies specific to Physical Arrangement as well as competencies specific to Literature Experience and Sense Perception training. In this manner, each of the treatment groups could serve as quasi- control groups for one another.

In addition to rating of classroom behavior "work samples", teacher performance was measured with a pre- posttraining paper-pencil assessment instrument with separate subscales measuring skills and knowledge involved in each of the three modules.

Upon completion of observation, written assessments, and training, each teacher was asked to rate her/his trainer's effectiveness and the training process specific to the particular competency module on which she/he was trained.

Operational Field Test Measurement

Instrumentation

Two methods, a classroom performance "work sample" rating and a paper-pencil exam, were used to assess pre-post, experimental-control group differences in teacher performance. Both methods were designed to measure key competencies in each of the three learning modules and paralleled one another in content subscales. The methods represented different ends on a verbal-behavioral performance continuum. That is, the paper-pencil method required trainees to utilize written test taking skills while the work sample method required trainees to behaviorally demonstrate competencies in a classroom situation. Copies of the work sample and paper-pencil instruments are included in Appendix A.

Paper-Pencil Method

The paper-pencil exam consisted of three sections, each covering competencies within one of the three training modules (Table 6). The exam contained 40 items totalling 85 points and took approximately 30-45 minutes to complete (there was no time limit specified). The test was administered by local trainers to all 32 teachers as a single instrument at the first and last workshop. Each workshop trainer was responsible for administering the test to all trainees enrolled in her/his workshop. All multiple choice items were scored at the Laboratory according to a

master key. The open-ended items were scored by two raters who had previously been trained to an 85% level of agreement reported in the reliability section.

Work Sample Method

The primary objective of the work sample method was to obtain measures of teachers' classroom application of competencies taught in all three training workshops. The work sample data was collected by a staff of ten part-time research assistants. Trained on procedures and equipped with cameras, tape recorders, books, sensory materials, and a variety of forms and written procedures, individual research assistants made four one-day classroom visits to each of the teachers in the study sample. The first two visits took place during the two weeks immediately before training; the second two visits took place during the two weeks immediately after training. The research assistants returned with records of the physical arrangement of each classroom and with audiotape recordings of each teacher interacting with children in her/his own classroom during the conduct of story-reading and sense perception learning experiences. These "work samples" were rated at the Laboratory according to how well they met the competency criteria of each unit. A more detailed description of the work sample method and procedures accompanies the work sample instruments in Appendix A.

Stimulus Materials

In order to standardize the classroom situations in which work sample date were collected, specific sets of materials and stimulus conditions were developed. They were designed to allow teachers to

Table 6. Distribution of possible points within the paper-pencil measure subscales.

Section	Competency Subscales	Possible Score
Physical Arrangement (PA)	Balance	15*
	Quality	14*
	Integration	11*
	TOTAL	30
Literature Experiences (LE)	Diversity	14
	Selection	6*
	Personalization	3*
	TOTAL	21
Sense Perception (SP)	Diversity	14*
	Awareness	14
	Diagnostics	14*
	TOTAL	33

*Subscales containing items included in other subscales.

produce particular training-related competencies. The various sets of stimulus materials and conditions are described below.

Physical Arrangement of the Classroom. A research assistant accompanied each teacher on a tour of the classroom. The assistant diagrammed the classroom's physical layout and interviewed the teacher regarding the purpose, use, and arrangement of each activity area within the room. The tour proceeded in a clockwise direction, and several photographs of each area were taken. The data was then brought to the laboratory for scoring. The performance variables scored were balance of the frequency of area types within a classroom, quality of the internal structure of individual areas, and integration of area types within the classroom.

Literature Experiences for Young Children. Each teacher conducted a story telling session with a group of 3-5 children using a children's book sent to her one week earlier. One half of the 32 teachers used the book Snowy Day on the first and third observation occasions and the book Birthday Presents on the second and fourth occasions. The remaining 16 teachers used the books in reverse order. The session was tape recorded and brought to the laboratory for scoring of diversity of teaching strategies used in presenting a story to children, and personalization of the literature experience for the individual children involved in it. In addition, a set of children's books was presented to each teacher to rate his/her selection of books for using with young children.

Developing Children's Senses and Perceptions. A set of beads of different colors and shapes and a set of 6 small containers of spices were assembled for teachers to use in preparing and conducting sensory experiences. One half of the teachers used beads for the sensory experiences on the first and third occasions and spices on the second and

fourth occasions. The remaining half of the teachers used the materials in reverse order: spices on first and third occasions, beads on second and fourth. The teachers conducted the experiences with a group of 3-5 children different from those involved in the story reading activities. The experiences were tape recorded and the tapes brought to the laboratory for scoring. The performance variables scored were awareness of multiple sensory uses of materials, diversity of teaching strategies used in conducting a sensory experience with young children and diagnosis of children's level of involvement, preferred sense mode, and cognitive skills used during sensory experiences.

Defining Teaching Competencies

As stated earlier, the three modules selected as experimental treatments for the OFT study (Physical Arrangement of the Classroom; Enriching Literature Experiences of Children; Developing Children's Sense Perception), were chosen because of their relative independence from one another in terms of content and because they sampled important teaching skill and content domains in early childhood education. To achieve the objectives of the OFT study, it was necessary to operationally define the primary instructional objectives of each OFT module as observable teaching behaviors. This was accomplished by specifying the conceptual organization of the FLS and its relationship with various early childhood teaching domains and teaching competencies. The resulting hierarchical model is presented in Figure 2.

At the top of the hierarchy are the three major Cluster Domains[2]

[2]The Cluster Domains were not originally a part of the FLS. They were conceived during the development of FLS modules to provide a means of organizaing groups of modules with similar goals and objectives. Thus, modules within them vary in their "degree of fit".

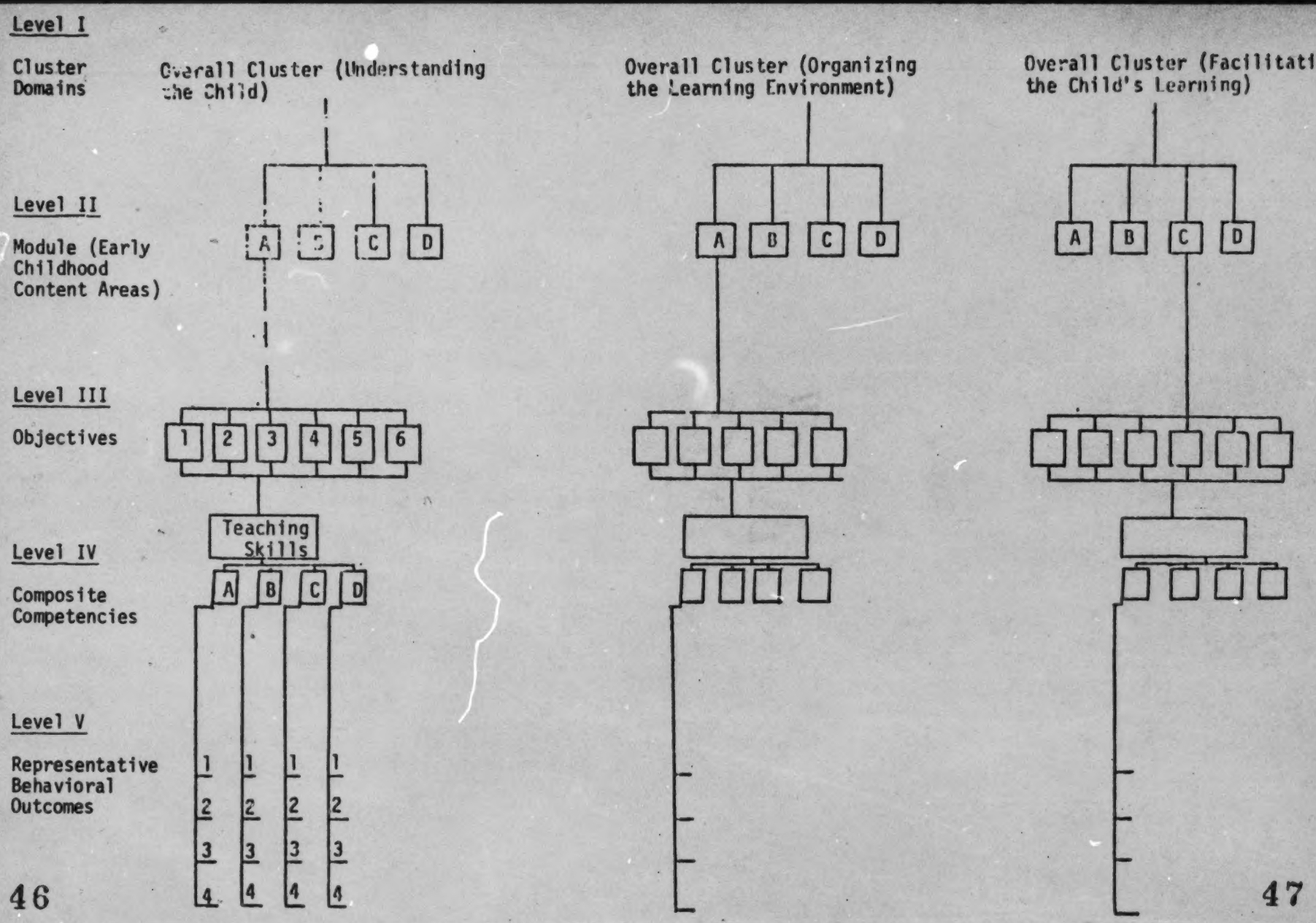


Figure 2. A conceptual hierarchy of the Flexible Learning System

(Level I). These represent broad teaching goals within the field of early childhood education. Within each domain lies a set of competency modules with more or less similar teaching goals and objectives (Level II). Each module represents a particular early childhood content area. Beneath each module lies a set of explicitly stated learning objectives (Level III). These objectives serve to focus the trainee's attention on the scope of the content of a module. They are usually interrelated and sequential, as they reflect the learning process embedded in the module. In order to achieve them, trainees must possess or acquire several teaching competencies (Level IV) which are composed of combinations of interrelated teaching behaviors (Level V). Thus, these FLS competencies are considered to be composites of a set of specific teaching skills (Level IV) necessary for mastering FLS goals and objectives. For example, three major competencies were defined for the module Physical Arrangement of the Classroom: (a) balance, (b) quality, and (c) integration. These represent several teaching skills provided via the module's training process, e.g., in order for a teacher to display the competency of quality of classroom areas, she/he must be able to plan learning activities for children, conceptualize spatial arrangement of areas, secure activity-related materials, etc. Thus, the conceptual hierarchy of the FLS suggests the relationship of the specific operational measures of competencies assessed in the research study to the broad goals of the system as a whole.

Establishing such a hierarchy is helpful in identifying levels of competency assessment which achieve an appropriate balance between abstract teaching goals and micro-teaching behaviors. If the competencies reflect only broad teaching goals, they are usually too vague to accurately define

a particular teaching skill and the task of reliably observing them can become nearly impossible; on the other hand, if competencies are over-refined, the essence or art of skilled teaching becomes fragmented into innumerable bits of behavior difficult to translate into competencies.

To establish a workable balance between over-and underrefinement of competencies for the OFT study, each of the three modules underwent rigorous content analysis. A brief description of each module's instructional purpose and design are presented below. Intensive discussions between the authors and evaluators focused the development of instruments for defining behavioral indicators of module competencies.

Physical Arrangement of the Classroom

This module assists teachers and teaching assistants in achieving an optimal arrangement of the classroom working spaces and materials for accomodating the wide range of classroom activities and child-groupings often found in "open-education" classrooms. The module is divided into three sections. The first section acquaints the learner with six types of classroom areas and with criteria for their adequacy. Criteria for matching arrangement of areas to the functional demands of the children's activities are also presented. Attention is devoted to issues that need to be considered in eliminating problems that typically occur in open rooms. The second section deals with types of overall classroom arrangements. The third section discusses a variety of inexpensive resources that can be used to create desired arrangements. Each section includes a number of activity worksheets, each with a follow-up to help the learner integrate her/his performance in terms of the content of the module. The instructional objectives and associated competencies for this module are presented in

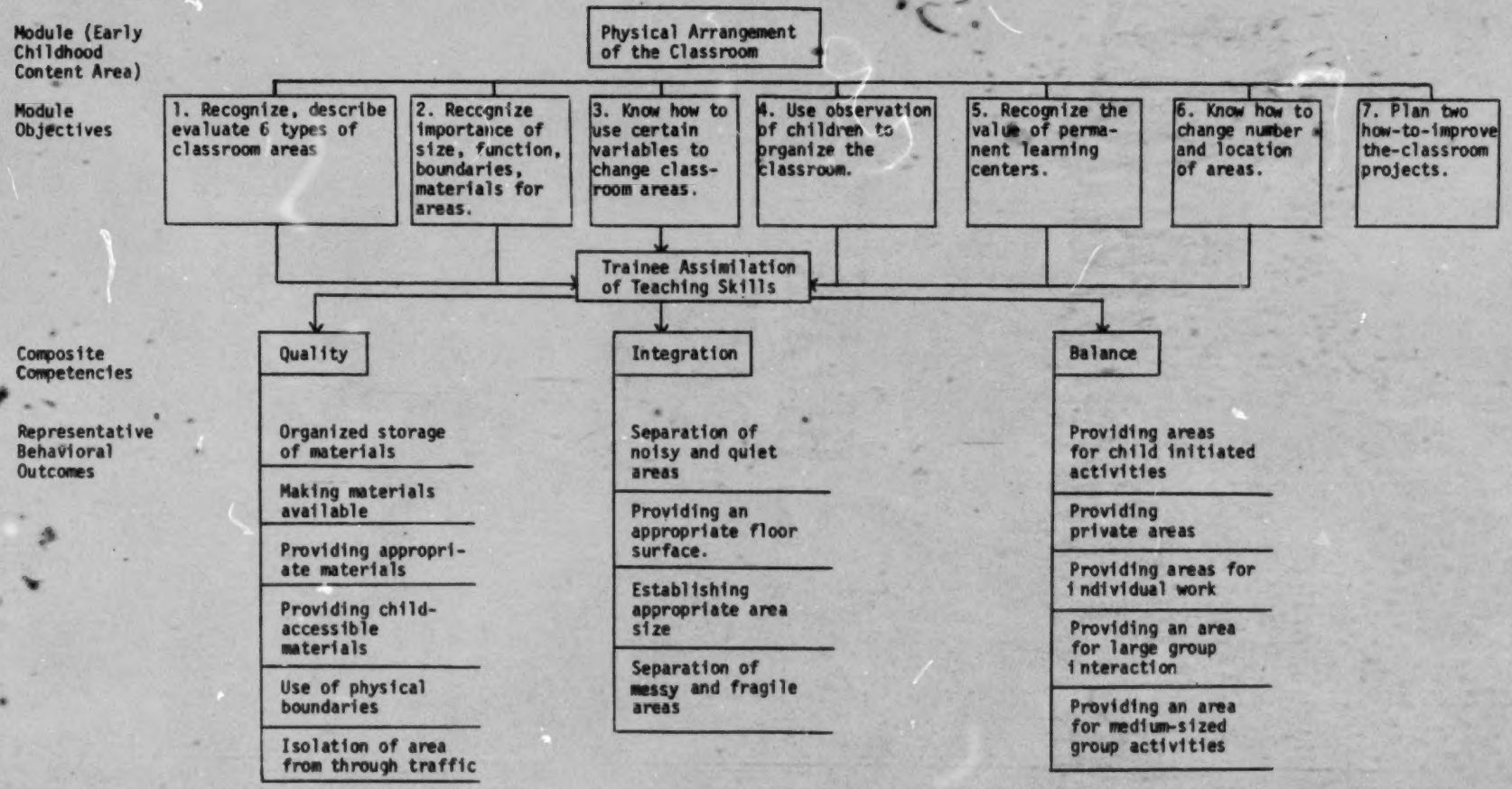

Module (Early Childhood Content Area)
Physical Arrangement of the Classroom
Module Objectives
1. Recognize, describe evaluate 6 types of classroom areas
2. Reccgnize importance of size, function, boundaries, materials for areas.
3. Know how to use certain variables to change classroom areas.
4. Use observation of children to organize the classroom.
5. Recognize the value of permanent learning centers.
6. Know how to change number and location of areas.
7. Plan two how-to-improve the-classroom projects.
Trainee Assimilation of Teaching Skills
Composite Competencies
Representative Behavioral Outcomes
Quality
Organized storage of materials
Making materials available
Providing appropriate materials
Providing child-accessible materials
Use of physical boundaries
Isolation of area from through traffic
Integration
Separation of noisy and quiet areas
Providing an appropriate floor surface.
Establishing appropriate area size
Separation of messy and fragile areas
Balance
Providing areas for child initiated activities
Providing private areas
Providing areas for individual work
Providing an area for large group interaction
Providing an area for medium-sized group activities

Figure 3. Competency hierarchy for the module Physical Arrangement of the Classroom

Figure 3.

Enriching Literature Experiences of Children

This learning module focuses on using children's literature. It develops competency in planning story reading sessions. The module consists of an introduction and three content sections. The introduction includes activities which help the trainee to recall her/his past experiences with children's literature, as well as providing the trainee with personal literature experiences. Section one helps the trainee become increasingly familiar with children's literature, establish workable guidelines for selection of literature, and plan the story session environment. Section two devotes attention to the reasons for using literature, the possibilities of literature, and ways to use literature to fulfill its potential for self-enchancement in children. Section three concludes the module by integrating what was presented in each of the preceeding sections. Each section contains a number of paper-pencil tasks and other activities, and includes a follow-up to help the learner integrate her/his performance with the module content. Figure 4 presents the objectives and competencies embedded in the module.

Developing Children's Sense Perception

This learning module consists of written materials on sensory and perceptual development along with a number of child-related activities that the learner pursues in class during the training program. These activities are directed toward acquainting the trainee with specific concepts and skills related to development of children's sense and perception. Guidelines for child observation and use of observation data are provided along with knowledge tests for learner self-assessment. The objectives and competencies for this module are presented in Figure 5.

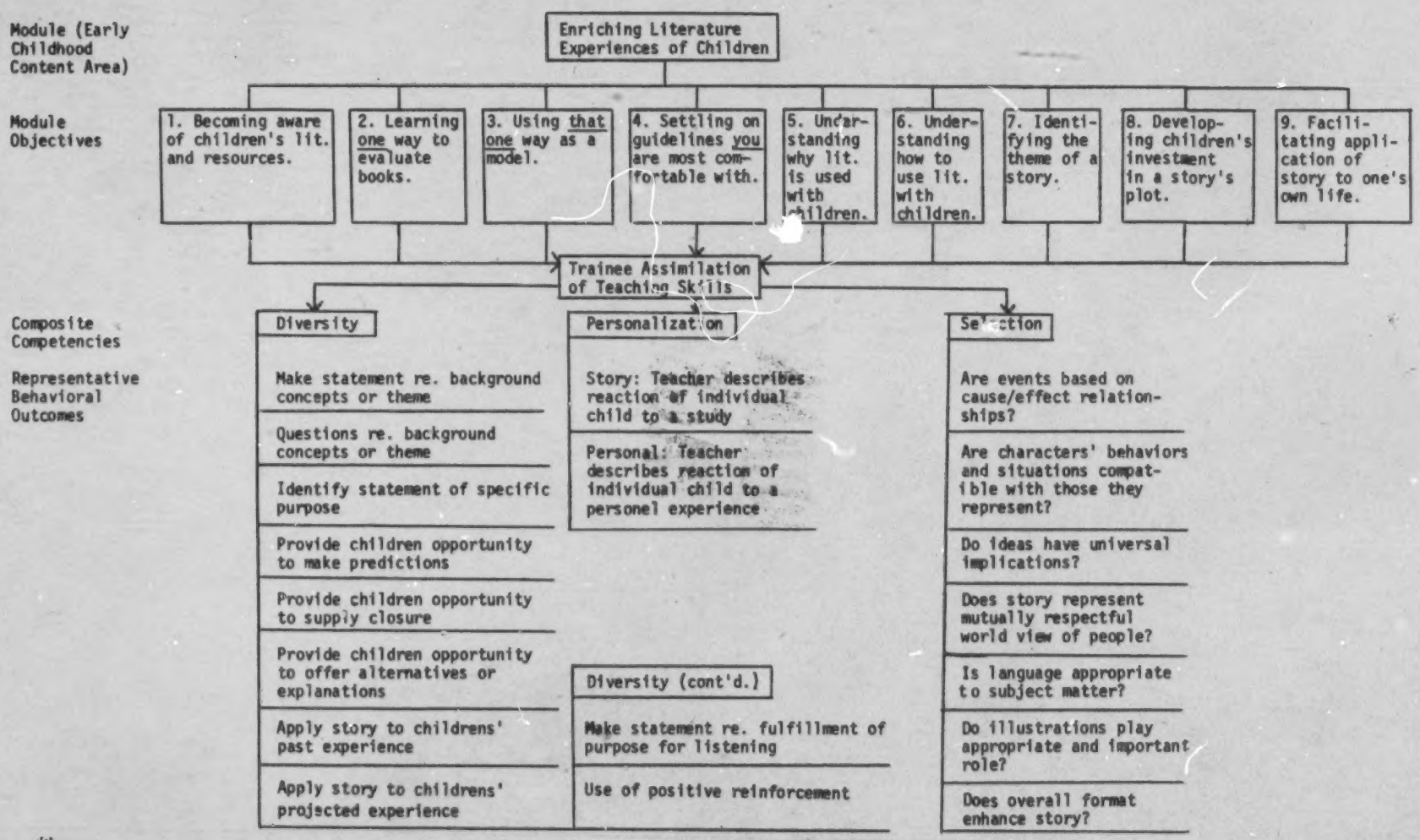

Module (Early Childhood Content Area)
Enriching Literature Experiences of Children
Module Objectives
1. Becoming aware of children's lit. and resources.
2. Learning one way to evaluate books.
3. Using that one way as a model.
4. Settling on guidelines you are most comfortable with.
5. Understanding why lit. is used with children.
6. Understanding how to use lit. with children.
7. Identifying the theme of a story.
8. Developing children's investment in a story's plot.
9. Facilitating application of story to one's own life.
Trainee Assimilation of Teaching Skills
Composite Competencies
Representative Behavioral Outcomes
Diversity
Make statement re. background concepts or theme
Questions re. background concepts or theme
Identify statement of specific purpose
Provide children opportunity to make predictions
Provide children opportunity to supply closure
Provide children opportunity to offer alternatives or explanations
Apply story to childrens' past experience
Apply story to childrens' projected experience
Personalization
Story: Teacher describes reaction of individual child to a study
Personal: Teacher describes reaction of individual child to a personel experience
Diversity (cont'd.)
Make statement re. fulfillment of purpose for listening
Use of positive reinforcement
Selection
Are events based on cause/effect relationships?
Are characters' behaviors and situations compatible with those they represent?
Do ideas have universal implications?
Does story represent mutually respectful world view of people?
Is language appropriate to subject matter?
Do illustrations play appropriate and important role?
Does overall format enhance story?

Figure 4. Competency hierarchy for the module Enriching Literature Experiences of Children

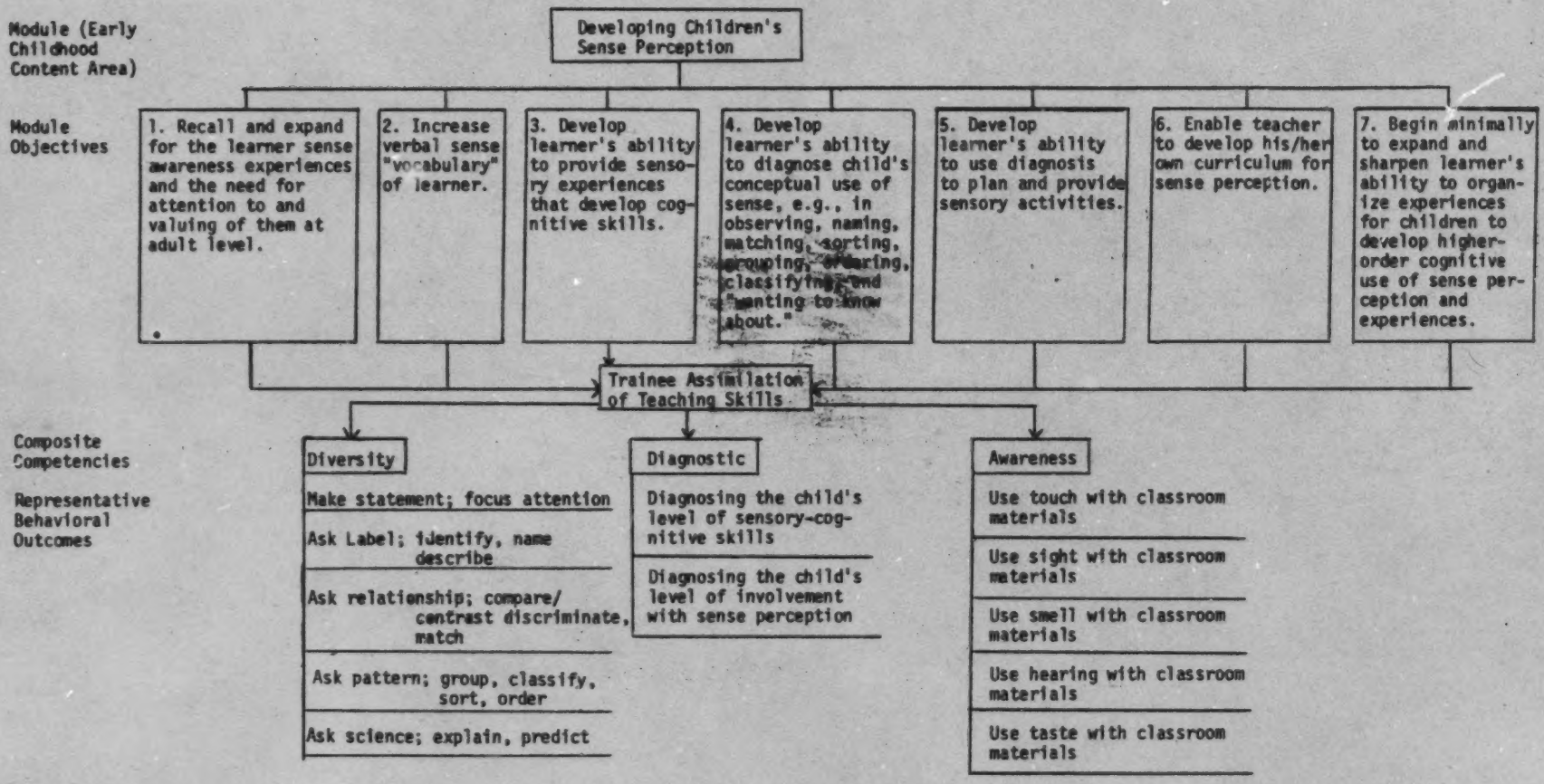

Module (Early Childhood Content Area)
Developing Children's Sense Perception
Module Objectives
1. Recall and expand for the learner sense awareness experiences and the need for attention to and valuing of them at adult level.
2. Increase verbal sense "vocabulary" of learner.
3. Develop learner's ability to provide sensory experiences that develop cognitive skills.
4. Develop learner's ability to diagnose child's conceptual use of sense, e.g., in observing, naming, matching, sorting, grouping, ordering, classifying, and "wanting to know about."
5. Develop learner's ability to use diagnosis to plan and provide sensory activities.
6. Enable teacher to develop his/her own curriculum for sense perception.
7. Begin minimally to expand and sharpen learner's ability to organize experiences for children to develop higher-order cognitive use of sense perception and experiences.
Trainee Assimilation of Teaching Skills
Composite Competencies
Representative Behavioral Outcomes
Diversity
Make statement; focus attention
Ask Label; identify, name describe
Ask relationship; compare/ contrast discriminate, match
Ask pattern; group, classify, sort, order
Ask science; explain, predict
Diagnostic
Diagnosing the child's level of sensory-cognitive skills
Diagnosing the child's level of involvement with sense perception
Awareness
Use touch with classroom materials
Use sight with classroom materials
Use smell with classroom materials
Use hearing with classroom materials
Use taste with classroom materials

Figure 5. Competency hierarchy for the module Developing Children's Sense Perception

Operational Field Test Results

The two research questions addressed in this study were: (a) whether the performance of teachers who receive FLS training in a specific area improves in relation to other teachers receiving FLS training in other areas or no training at all, and (b) whether teachers with different educational backgrounds perform differentially on paper-pencil versus classroom measures of skills.

Several analyses were conducted on the pre- and posttraining data to answer the first question of FLS training effects on performance. First, an analysis-of-variance was carried out to detect pretraining differences among the four groups of teachers. In effect, this analysis constituted a check on the effectiveness of the random assignment of teachers to treatment conditions. Second, each treatment group's change in pre- to posttraining performance was analyzed using a group by trials analysis-of-variance. Third, a series of multiple t-test comparisons of posttraining performance between treatment and quasi-control (i.e., other FLS treatment groups receiving training in different areas combined with a control group receiving no training) groups were carried out. Fourth, t-test posttraining comparisons were made between each treatment group versus each other treatment group in the study. For example, the Physical Arrangement treatment group was compared separately to the Sense Perception treatment group, the Literature Experience treatment group, and the Control group.

For the second research question -- effects of educational background on performance measured with different methods -- a similar series of analyses were run. First, an analysis-of-variance was conducted to determine pretraining performance differences between the low education

(high school diploma or less) and high education groups (B.A. degree or more). Second, posttraining performance of low education teachers within a treatment was compared with high education teachers within the same treatment using multiple t-tests. Third, comparisons were made between all low education teachers who had received training and all high education teachers who had received training. In each comparison, the t-value was based on unbiased pooled estimates of variance unless there were significant differences among group variances ($p < .05$, in which case separate estimates of variance were used to compute t-values).

It is important to note that reliability for the work sample measures was determined by generating coefficients of generalizability (Cronbach, et al, 1972) for each measure. These coefficients and test-retest reliability for the paper-pencil exam are reported in Appendix B. The results of the generalizability analysis were the basis for defining parameters of dependent competencies to be used in the data analyses conducted to answer the two primary research questions. The generalizability results indicated that the dependent competency diversity specific to the Literature Experience treatment condition should be analyzed separately according to stimulus materials used to elicit it (i.e., the children's books Snowy Day or Birthday Presents). The Literature Experience dependent competency of personalization was analyzed separately according to the data collection during which it was recorded (occasion 1, 2, 3, or 4). The same was true for the Sense Perception dependent competencies of diversity and diagnostics -- performance on each differed according to the data collection occasion during which it was recorded.

In all of the posttraining analyses, a sample of 31 teachers served

as the unit of analysis due to the attrition of one teacher from the original sample.

The remainder of the results section will present findings according to results of:

(a) pretraining analysis-of-variance between treatment groups;

(b) pre- to posttraining analysis of variance of gains;

(c) multiple comparisons of combined quasi-, and pure controls to treatment groups;

(c) pre- and posttraining effects and differences between teachers with limited or extensive formal education.

Additional data on variables which may have mediated module training effectiveness (e.g., trainer effectiveness) and the methods for collecting the data are presented in Appendix C.

Treatment Pretraining Results

Table 7 presents the pretraining analysis-of-variance for all FLS training groups. Of the 25 dependent competencies, there were significant differences among groups on the Literature Experience competency of personalization as measured by the paper-pencil exam and the Sense Perception competency of awareness as measured by the paper-pencil exam. The Sense Perception treatment group scored relatively higher and the Control group relatively lower on personalization that the Physical Arrangement and Literature Experience groups, suggesting that the significant difference may be a result of the discrepancy between the Sense Perception and Control group's scores. The Control group scored substantially lower on the Sense Perception competency of awareness. In addition, the Control group scored somewhat lower on all of the paper-pencil total scores, though there were

Table 7. Pretraining performance means, standard deviations and F values for the OFT training groups.

Competencies			Treatment Group: Physical Arrangement (n=8) Pretraining Mean	s.d.	Literature Experience (n=8) Pretraining Mean	s.d.	Sense Perception (n=7) Pretraining Mean	s.d.	Control (n=8) Pretraining Mean	s.d.	Between Group F ratio
Phys. Arr.	Work Sample	Balance	4.25	0.38	3.81	0.75	4.43	0.61	4.31	0.65	1.50
		Quality	45.69	3.76	47.92	5.21	47.24	5.15	46.00	3.74	0.43
		Integration	6.81	1.46	6.13	1.79	5.36	1.14	6.06	1.61	1.13
	Paper-Pencil	Balance	9.88	2.17	9.13	2.85	9.86	1.86	7.63	3.42	1.23
		Quality	7.44	1.40	6.69	2.17	7.43	2.03	6.19	2.25	0.73
		Integration	6.25	1.91	6.25	1.31	6.43	1.54	5.63	2.49	0.27
		Total Score	17.56	2.99	16.56	3.23	17.43	3.21	14.19	4.76	1.46
Lit. Exps.	Work Sample	Selection	6.38	2.50	6.75	1.98	6.29	3.25	5.00	2.98	0.64
		Diversity-SD	6.00	2.51	5.63	2.67	5.43	2.07	5.63	2.62	0.07
		Diversity-BP	5.00	2.07	5.13	1.46	4.29	2.36	5.88	2.53	0.70
		Personalization-01	0.25	0.71	0.50	1.07	0.29	0.76	0.75	1.39	0.40
		Personalization-02	1.50	1.72	0.38	0.99	0.38	0.69	0.38	0.69	2.25
	Paper-Pencil	Selection	3.63	1.30	3.63	0.92	4.43	1.40	3.25	0.46	1.58
		Diversity	9.50	3.02	9.13	1.55	9.43	1.72	8.13	1.81	0.70
		Personalization	1.75	0.71	1.50	0.54	2.29	0.95	1.13	0.64	3.45*
		Total Score	13.63	4.37	13.38	2.26	14.29	1.98	11.63	1.85	1.24
Sense Perc.	Work Sample	Diversity-01	6.50	1.93	8.00	3.21	7.57	2.07	6.13	1.64	1.14
		Diversity-02	6.38	2.39	6.88	1.55	6.14	1.77	7.00	2.27	0.30
		Diagnostic-01	1.00	2.45	0.88	1.81	0.43	0.79	0.38	1.06	0.27
		Diagnostic-02	0.75	1.39	0.13	0.35	0.14	0.38	0.00	0.00	1.60
		Awareness	5.13	5.94	5.00	4.75	6.00	3.46	5.25	1.98	0.08
	Paper-Pencil	Diversity	8.19	3.29	9.31	1.89	9.36	2.02	8.44	2.35	0.46
		Diagnostic	7.88	3.44	8.81	1.83	8.86	2.32	8.13	1.55	0.33
		Awareness	4.34	1.45	5.56	1.76	5.07	1.90	2.97	2.19	2.97*
		Total Score	14.66	4.82	16.94	2.84	16.43	3.30	13.16	4.58	1.47

*p < .05
**p < .01

no significant differences. It should also be noted that the pretraining scores of each group on the Literature Experience personalization competency for occasion one were extremely low, while the occasion two scores were relatively higher but still fairly low. The same was true for Sense Perception diagnostics, low scores were evident for each group on occasions one and two.

Physical Arrangement Treatment Effects

The Physical Arrangement (PA) treatment group pre- to posttraining gains are shown in Table 8. The group performed significantly better after training on six of seven competencies: Work sample balance and quality paper-pencil balance, quality, integration, and total score.

Table 9 presents the results of comparisons between the PA group and the combined control group (i.e., all other teachers in the study) on PA competencies. The PA group performed significantly better on the work sample competencies of balance and quality and the paper-pencil competencies of balance, quality, integration, and total score.

When compared to each of the other treatment groups (Table 10), the PA group significantly outperformed the Literature Experience group on the work sample competency of quality, and the paper-pencil competencies of balance, quality, integration, and total score. The PA group significantly outscored the Sense Perception group on the paper-pencil total score and the Control group on work sample balance and quality, and paper-pencil integration. Overall, the PA group had higher mean scores than any other group on every PA measure.

Literature Experience Treatment Effects

Table 11 shows the pre- to posttraining performance changes for the

Table 8. Pre- to posttraining performance means, standard deviations, and F values for Physical Arrangement treatment group.

Physical Arrangement Competencies		Physical Arrangement Group (n=8) Pre-training Performance Mean	s.d.	Posttraining Performance Mean	s.d.	Repeated Measures F ratio
Work Sample	Balance	4.25	0.38	5.25	1.00	6.59*
	Quality	45.67	3.76	52.93	4.16	9.10*
	Integration	6.81	1.46	7.31	1.73	0.61
Paper-Pencil	Balance	9.88	2.17	12.25	2.37	6.89*
	Quality	7.44	1.40	8.75	2.07	5.22*
	Integration	6.25	1.91	8.81	1.69	19.64**
	Total Score	17.56	2.99	21.50	2.96	21.59**

* $p < .05$
** $p < .01$

Table 9. Posttraining performance means, standard deviations, and t values for the Physical Arrangement treatment and control groups.

Physical Arrangement Competencies		Treatment Group: Physical Arrangement (n=8) Mean	s.d.	Control Group (n=23) Mean	s.d.	t
Work Sample	Balance	5.25	1.00	4.50	0.77	2.20*
	Quality	52.93	4.16	47.82	4.38	2.88**
	Integration	7.31	1.73	6.52	1.78	1.09
Paper-Pencil	Balance	12.25	2.37	9.68	2.67	2.47*
	Quality	8.75	2.07	7.04	1.93	2.12*
	Integration	8.83	1.69	6.85	1.82	2.67**
	Total Score	21.50	2.96	17.04	3.75	3.03**

*p <.05
**p <.01

Table 10. Posttraining performance means, standard deviations, and t values for the Physical Arrangement treatment group vs. each other group.

Physical Arrangement Competencies		Physical Arrangement (n=8)		Literature Experiences (n=8)			Sense Perception (n=7)			Control (n=8)		
		Mean	s.d.	Mean	s.d.	t	Mean	s.d.	t	Mean	s.d.	t
Work Sample	Balance	5.25	1.00	4.81	0.92	0.91	4.43	0.45	2.09△	4.25	0.80	2.21
	Quality	52.93	4.16	48.11	4.09	2.34	48.18	5.24	1.96	47.20	4.40	2.67*
	Integration	7.31	1.73	7.06	1.35	0.32	5.50	1.68	2.05	6.88	2.05	0.46
Paper-Pencil	Balance	12.25	2.38	8.75	1.83	3.30**	10.14	3.18	1.47	10.00	3.02	1.66
	Quality	8.75	2.07	7.00	1.10	2.11*	7.36	1.38	1.51	6.81	2.95	1.52
	Integration	8.81	1.69	6.63	1.03	3.13**	7.29	2.36	1.46	6.69	2.09	2.24*
	Total Score	21.50	2.96	16.00	2.32	4.14**	17.93	3.25	2.23*	17.31	5.27	1.96

* $p < .05$
** $p < .01$

△ t values based on variance estimates for separate groups.

Literature Experience (LE) treatment group. Mean scores were higher following training for six of the ten competencies, four of the means were appreciably higher ($F \geq 1.0$),[3] including the change on diversity - BP which reached significance.

The LE group had mean scores that were appreciably higher ($t \geq 1.0$) than the combined control group on the LE work sample measures of personalization - 03 (occasion 3), diversity - BP (Table 12); however, neither of the LE group scores were significantly higher. The LE group paper-pencil mean scores were not appreciably different than the combined control mean scores.

Compared to each of the other three groups (Table 13) the LE group means were higher on three of the work sample competencies; diversity - SD, diversity - BP, personalization - 03. On the work sample measure of the selection skill the LE group scored higher than both the SP and the Control groups. On the competencies measured by paper and pencil tests the LE group had higher mean scores than the Control group for selection, diversity, and the total score. In a total of 27 mean score comparisons the LE group had seven appreciably higher, and four noticeably lower ($t \geq 1.0$) than the mean scores of each of the three other groups. None of the mean score differences reached significance.

Sense Perception Treatment Effects

Table 14 shows the pre- to posttraining changes in the Sense Perception (SP) treatment group's performance. Average Performance scores improved noticeably following training on eight of the nine competencies. The

[3]To examine trends among mean performance scores, F and t values equal to or greater than 1.0 were designated as indicating appreciable or noticable differences between means. The differences that reached significance were specified.

Table 11. Pre- to posttraining performance means, standard deviations, and F values for Literature Awareness treatment group.

	Literature Awareness Competencies	Literature Awareness Group (n=8) Pre-Training Performance Mean	s.d.	Posttraining Performance Mean	s.d.	Repeated Measures F Ratio
Work Sample	Selection	6.75	1.98	6.00	1.69	0.56
	Diversity-BP	5.13	1.46	6.88	2.29	7.30*
	Diversity-SD	5.63	2.67	6.13	2.59	2.33
	Personalization 01-03	0.50	1.07	1.00	0.93	1.00
	Personalization 01-04	0.50	1.07	0.50	0.53	0.00
	Personalization 02-03	0.38	0.99	1.00	0.93	0.73
	Personalization 02-04	0.38	0.99	0.50	0.53	0.14
Paper-Pencil	Selection	3.63	0.92	4.00	1.41	0.80
	Personalization	1.50	0.54	1.75	1.04	1.00
	Diversity-LE	9.13	1.55	8.38	1.77	1.47
	Total Score	13.38	2.26	12.75	2.92	0.50

* $p < .05$

Note. F ratios were computed between each of the two pre-training (01-02) and the two posttraining (03-04) observations on the work sample measure, Personalization. The F ratios were not significant.

Table 12. Posttraining performance means, standard deviations, and t values for Literature Awareness treatment and control groups.

Literature Awareness Competencies		Treatment Group: Literature Awareness (n=8) Mean	s.d.	Control Group (n=23) Mean	s.d.	t
Work Sample	Selection	6.00	1.69	5.43	2.25	0.65
	Personalization-03	1.00	0.93	0.39	0.78	1.81
	Personalization-04	0.50	0.53	0.87	1.36	0.74
	Diversity-SD	6.13	2.59	5.35	2.08	0.86
	Diversity-BP	6.88	2.29	5.35	2.62	1.56
Paper-Pencil	Selection	4.00	1.41	3.96	1.26	0.08
	Personalization	1.75	1.04	1.87	0.63	0.39
	Diversity-LE	8.38	1.77	9.09	2.48	0.74
	Total Score	12.75	2.92	13.30	3.04	0.45

*$p < .05$
**$p < .01$

Table 13. Posttraining performance means, standard deviations, and t values for the Literature Experience treatment group vs. each other group.

Literature Experience Competencies		Literature Experience (n=8)		Physical Arrangement (n=8)			Sense Perception (n=7)			Control (n=8)		
		Mean	s.d.	Mean	s.d.	t	Mean	s.d.	t	Mean	s.d.	t
Work Sample	Selection	6.00	1.69	6.00	2.07	0.00	4.86	2.41	1.07	5.38	2.45	0.59
	Diversity-SD	6.13	2.59	5.88	2.59	0.19	5.00	1.41	1.02	5.13	2.17	0.84
	Diversity-BP	6.88	2.30	5.25	2.38	1.39	5.14	2.85	1.30	5.63	2.97	0.94
	Personalization-03	1.00	0.93	0.38	0.74	1.49	0.29	0.76	1.62	0.50	0.93	1.08
	Personalization-04	0.50	0.54	0.50	0.76	0.00	1.14	1.68	0.97Δ	1.00	1.60	0.84Δ
Paper-Pencil	Selection	4.00	1.41	4.38	0.74	0.66	4.29	0.95	0.45	3.25	1.67	0.97
	Diversity	8.38	1.77	10.25	2.66	1.66	9.57	1.99	1.23	7.50	2.07	0.91
	Personalization	1.75	1.04	2.00	0.54	0.61Δ	1.86	0.69	0.23	1.75	0.71	0.00
	Total Score	12.75	2.92	14.88	2.80	1.49	14.14	1.77	1.10	11.00	2.98	1.19

*p < .05
**p < .01

Δ t values based on variance estimates for separate groups.

Table 14. Pre- to posttraining performance means, standard deviations, and F values for Sense Perception treatment group.

Sense Perception Competencies		Sense Perception Group (n=7) Pretraining Performance Mean	s.d.	Posttraining Performance Mean	s.d.	Repeated Measures F ratio
Work Sample	Diversity 01-03	7.57	2.07	7.86	3.08	0.41
	Diversity 01-04	7.57	2.07	8.43	2.82	0.42
	Diversity 02-03	6.14	1.77	7.86	3.08	1.52
	Diversity 02-04	6.14	1.77	8.43	2.82	4.27
	Diagnostic 01-03	0.43	0.79	1.00	2.24	0.04
	Diagnostic 01-04	0.43	0.79	0.86	1.57	0.42
	Diagnostic 02-03	0.14	0.38	1.00	2.24	1.48
	Diagnostic 02-04	0.14	0.38	0.86	1.57	2.27
	Awareness	6.00	3.45	4.29	4.15	1.45
Paper-Pencil	Diversity	9.36	2.02	11.71	1.38	3.26
	Diagnostic	8.86	2.32	11.71	0.57	9.58*
	Awareness	5.07	1.90	6.46	2.76	8.94*
	Total Score	16.43	3.30	20.75	3.07	8.68*

* $p < .05$

Note. F ratios were computed between each of the two pre-training (01-02) and posttraining (03-04) observations on the work sample measures Diagnostic and Diversity. The F ratios were not significant.

improvement was significant on the paper-pencil competencies of diversity, diagnostics and total score.

The SP group had higher mean scores than the combined control group on four of the five SP work sample competencies two of these were appreciably higher and each of the four paper-pencil competencies (Table 15). Of these, the SP group scores on paper-pencil diagnostics awareness, and total score were significantly higher than the combined control group scores.

On all of the paper and pencil measures of SP skills the SP group was higher than each of the other groups (Table 16). This difference reached significance on the total score with the SP group higher than both the LE and Control groups. On the awareness and diversity competencies the SP was significantly higher than the Control group. Of the 15 mean score comparisons on the work sample measures of SP competencies the SP group was appreciably higher on seven and lower on two of the mean scores of each of the other three groups. For both diversity - 03 and diversity - 04 the SP group scored higher than any other group. None of these differences reached significance.

Educational Pre- and Posttraining Results

The pretraining analysis-of-variance results for the comparison of the high (n=16) and low (n=15) educational groups revealed eight significant differences in performance between the two groups (Table 17). The high education group significantly outperformed the low education group on five paper-pencil competencies: Literature Experiences diversity and total score; and Sense Perception diversity, diagnostic, and total score -- and two work sample competencies - the Literature Experience competency of diversity - SD (Snowy Day book) and the Sense Perception competency of diversity-01

Table 15. Posttraining performance means, standard deviations, and t values for the Sensory Perception treatment and control groups.

Sensory Perception Competencies		Treatment Group: Sensory Perception (n=7) Mean	s.d.	Control Group (n=24) Mean	s.d.	t
Work Sample	Diagnostic-03	1.00	2.24	0.54	1.47	0.64
	Diagnostic-04	0.86	1.57	0.63	1.74	0.32
	Awareness	4.29	4.15	6.75	3.86	1.46
	Diversity-03	7.86	3.08	6.50	2.21	1.31
	Diversity-04	8.43	2.82	6.17	2.62	1.98
Paper-Pencil	Diversity	11.71	1.38	9.96	2.64	1.68
	Diagnostic	11.71	0.57	9.85	2.63	3.22**
	Awareness	6.46	2.76	4.23	1.92	2.46*
	Total Score	20.75	3.07	16.29	4.42	2.48*

*$p < .05$
**$p < .01$

Table 16. Posttraining performance means, standard deviations and t values for the Sense Perception treatment group vs. each other group.

Sense Perception Competencies		Sense Perception (n=7) Mean	s.d.	Physical Arrangement (n=8) Mean	s.d.	t	Literature Experiences (n=8) Mean	s.d.	t	Control (n=8) Mean	s.d.	t
Work Sample	Diagnostic-03	1.00	2.24	1.00	2.14	0.00	0.13	0.35	1.02 Δ	0.50	1.41	0.53
	Diagnostic-04	0.86	1.57	1.63	2.77	0.65	0.00	0.00	1.55	0.25	0.71	0.94Δ
	Awareness	4.29	4.15	7.00	4.41	1.22	7.88	4.36	1.63	5.38	2.62	0.62
	Diversity-03	7.86	3.08	6.13	1.36	1.38Δ	7.63	3.11	0.14	5.75	1.49	1.65Δ
	Diversity-04	8.43	2.82	5.75	2.25	2.05	6.50	3.21	1.23	6.25	2.61	1.56
Paper-Pencil	Diagnostic	11.71	1.38	10.94	1.84	0.91	10.06	1.08	2.60	8.88	3.99	1.89Δ
	Awareness	6.46	2.76	5.25	1.20	1.13	4.56	1.34	1.66Δ	2.88	2.33	2.74*
	Diversity	11.71	0.57	10.05	2.45	1.36Δ	10.38	1.58	2.12	8.69	3.44	2.45*Δ
	Total Score	20.75	3.07	18.31	3.71	1.37	17.00	2.72	2.51*	13.56	5.42	3.09**

* $p < .05$
** $p < .01$

Δ t value based on variance estimates for separate groups.

Table 17. Pretraining performance means, standard deviations, and F values between two OFT training groups differing in educational background.

	Competencies		Low Education (n=15) Mean	s.d.	High Education (n=16) Mean	s.d.	Between Group F ratio
Work Sample	Phys. Arr.	Balance	4.07	0.32	4.31	0.81	1.19
		Quality	47.02	3.93	46.40	4.86	0.15
		Integration	5.90	1.54	6.31	1.57	0.55
	Lit. Exps.	Selection	5.60	2.61	6.56	2.68	1.02
		Diversity-SD	4.00	1.56	6.13	2.06	10.37**
		Diversity-BP	5.27	2.49	6.06	2.27	0.87
		Personalization-01	0.60	0.99	0.31	1.01	0.64
		Personalization-02	0.60	0.95	0.31	1.07	0.64
	Sense Perception	Diversity-01	6.20	1.94	7.80	2.29	4.11*
		Diversity-02	6.93	2.28	6.31	1.62	0.77
		Diagnostic-01	0.73	1.88	0.63	1.40	0.03
		Diagnostic-02	0.47	1.06	0.62	0.25	2.20
		Awareness	6.00	5.28	4.69	2.65	0.78
Paper Pencil	Phys. Arr.	Balance	10.07	2.66	8.19	2.49	4.14*
		Quality	6.50	2.18	7.31	1.71	1.34
		Integration	6.60	1.79	5.69	1.77	2.04
		Total Score	16.90	4.12	15.94	3.35	0.53
	Lit. Exps.	Selection	3.47	1.13	3.94	1.06	1.44
		Diversity	7.93	2.02	10.06	1.61	10.62**
		Personalization	1.60	0.83	1.69	0.79	0.09
		Total Score	11.80	3.05	14.50	2.00	8.60**
	Sense Perception	Diversity	7.60	2.32	9.94	1.91	9.43**
		Diagnostic	7.03	2.31	9.69	1.46	14.83**
		Awareness	3.90	2.22	5.00	1.70	2.42
		Total Score	13.13	4.33	17.25	2.68	10.29**

*p < .05
**p < .01

(occasion one). Conversely, the low education group significantly outscored the high education group on only the Physical Arrangement paper-pencil competency of balance. There were 15 appreciably ($t \geq 1.0$) different mean scores comparisons between the two groups. The high education group outscored the low education group on five work sample skills and eight paper-pencil measures. The low education teachers were appreciably higher on only two paper-pencil skills.

The t-tests between the scores of the high (n=12) and low (n=11) educational groups who had reveived FLS training are shown in Table 18. The high education group significantly outscored the low education group on the work sample skills - PA balance and LE diversity - SD -- and on the paper-pencil skill SP diversity. Of the 14 mean scores appreciably different the low education teachers outscored the high education teachers on four work sample competencies while the high outscored the low on three. The results were reversed for the paper-pencil performance differences; high education teachers outscored low education teachers on four competencies, the low outscored the high on three.

It is informative to note that the low education group had greater score variance (standard deviations) than the high education group on 18 of the pre- and 17 of the posttraining measures of the 25 competencies, suggesting greater variability in performance of the low education sample as compared to the high education group.

Table 19 presents the comparisons of PA low and high education teachers an PA competencies. The low education group scored significantly better than the high education group on paper-pencil integration, whereas the high education group significantly outperformed the low education group on

Table 18. Posttraining performance means, standard deviations, and t values for high education vs. low education groups that received training.

	Competencies		Low Education (n=11) Mean	s.d.	High Education (n=12) Mean	s.d.	t
Work Sample	Physical Arrangement	Balance	4.41	0.77	5.25	0.78	2.59*
		Quality	49.79	4.50	49.83	5.39	0.02
		Integration	6.59	1.96	6.75	1.54	0.22
	Literature Experience	Selection	5.82	2.18	5.50	1.98	0.37
		Diversity-SD	4.64	2.20	6.67	1.87	2.39*
		Diversity-BP	5.36	2.62	6.17	2.48	0.76
		Personalization-03	0.91	0.94	0.25	0.62	1.99
		Personalization-04	0.82	1.17	0.58	1.00	0.52
	Sense Perception	Diversity-03	7.00	1.79	7.33	3.28	0.30
		Diversity-04	7.45	3.11	6.25	2.67	1.00
		Diagnostic-03	0.64	1.80	0.75	1.76	0.15
		Diagnostic-04	0.64	1.57	1.00	2.22	0.45
		Awareness	7.91	4.89	5.17	3.59	1.54
Paper Pencil	Physical Arrangement	Balance	10.91	2.74	9.92	2.91	0.84
		Quality	7.23	1.56	8.17	1.76	1.35
		Integration	8.36	2.34	6.88	1.28	1.98
		Total Score	18.50	3.92	18.50	3.47	0.00
	Literature Experience	Selection	4.55	0.69	3.92	1.24	1.48
		Diversity	9.27	3.00	9.50	1.31	0.23
		Personalization	2.09	0.83	1.67	0.65	1.37
		Total Score	14.27	2.94	13.58	2.39	0.62
	Sense Perception	Diversity	10.14	1.55	11.54	1.29	2.37*
		Diagnostic	10.09	2.25	11.50	0.83	1.96
		Awareness	5.27	1.96	5.48	1.98	0.25
		Total Score	17.50	3.87	19.60	2.75	1.51

* $p < .05$

t value based on variance estimates for separate groups.

Table 19. Posttraining performance means, standard deviations, and t values for Physical Arrangement high versus low education teachers.

Physical Arrangement Competencies		Physical Arrangement: Low Education (n=4) Mean	s.d.	High Education (n=4) Mean	s.d.	t
Work Sample	Balance	4.75	0.96	5.75	0.87	1.55
	Quality	51.51	5.85	54.35	0.84	0.96
	Integration	8.13	1.97	6.50	1.15	1.42
Paper-Pencil	Balance	13.50	1.29	11.00	2.71	1.67
	Quality	7.38	2.06	10.13	0.85	2.47*
	Integration	10.25	0.50	7.38	0.95	5.37**
	Total Score	21.13	3.42	21.88	2.90	0.33

* $p < .05$
** $p < .01$

paper-pencil quality. The PA low education group evidenced greater score variability (standard deviations) on five of the seven competency scores. The PA low education group outscored the PA high education group three to two on the five appreciably different mean score comparisons.

The differences between high and low education teachers within each treatment group at posttraining were examined according to method-of-measurement. Table 20 presents the posttraining comparisons between LE low and high education teachers on LE competencies. The LE low education teachers had noticeably higher mean scores on the work sample competencies of selection and personalization - 04 and the paper-pencil competency of selection, though none of the differences achieved significance. The LE high education group was appreciably higher on the SP work sample skill diversity.

The scores of SP low and high education teachers are presented in Table 21. The low education teachers did significantly better on work sample diversity - 04 than the high education teachers and had appreciably higher mean scores on one work sample and one paper-pencil competency.

Summarizing the posttraining mean score differences of appreciable magnitude for each of three treatment groups the resulting tallys were four to two on the work sample skills and four to one on paper-pencil skills, both in favor of the low education teachers. The PA high education group twice scored appreciably higher than the low education group. One of those comparisons reached significance, the LE high education group had one mean score noticeably higher than the LE low education group. The SP high education group did not score noticeably higher than the SP low education group.

Table 20. Posttraining performance means, standard deviations, and t values for Literature Experience high versus low education teachers.

	Literature Experience Competencies	Literature Experiences: Low Education (n=4) Mean	s.d.	High Education (n=4) Mean	s.d.	t
Work Sample	Selection	6.75	0.50	5.25	2.22	1.32
	Diversity-SD	4.50	2.08	7.75	2.06	2.22
	Diversity-BP	6.25	2.36	7.50	2.38	0.75
	Personalization-03	1.25	0.96	0.75	0.96	0.74
	Personalization-04	0.75	0.50	0.25	0.50	1.41
Paper-Pencil	Selection	4.50	1.00	3.50	1.73	1.00
	Diversity	8.00	2.31	8.75	1.26	0.57
	Personalization	1.75	1.26	1.75	0.96	0.00
	Total Score	13.25	3.40	12.25	2.75	0.46

Table 21. Posttraining performance means, standard deviations, and t values for Sense Perception high versus low education teachers.

Sense Perception Competencies		Sense Perception: Low Education (n=3) Mean	s.d.	High Education (n=4) Mean	s.d.	t
Work Sample	Diversity-03	9.00	1.00	7.00	4.00	0.83
	Diversity-04	11.00	2.00	6.50	1.29	3.65*
	Diagnostic-03 Δ	0.33	0.58	1.50	3.00	0.65
	Diagnostic-04	0.67	1.15	1.00	2.00	0.25
	Awareness	6.00	5.57	3.00	2.94	0.94
Paper-Pencil	Diversity	11.33	1.53	12.00	1.41	0.60
	Diagnostic Δ	11.83	0.29	11.63	0.75	0.45
	Awareness	7.75	1.64	5.50	3.24	1.08
	Total Score	21.92	2.90	19.88	3.30	0.85

* $p < .05$

Δ t value based on variance estimates for separate groups.

Discussion of OFT Findings

In an attempt to address the two primary OFT questions of a) FLS training effects and b) differences in performance of teachers with relatively little or extensive formal education, this section includes:

(a) effects of OFT training on teacher performance;

(b) summary of OFT training effectiveness;

(c) effects of educational level and methods of measurement;

(d) summary of OFT high and low education teacher's performance;

(e) recommendations for the FLS.

Teacher training is a complex enterprise and adequate illumination of a particular training system's processes and effectiveness is best gained from several sources, thus, the statistical results from the OFT have been supplemented in this section by ethnographic data gleaned with various methods described in Appendix C.

Effects of OFT Training

There were two significant paper-pencil performance differences before training, and on both measures the Control group scored substantially lower than the other OFT groups. Accordingly, posttraining differences on these measures -- Literature Experience paper-pencil personalization and Sense Perception paper-pencil awareness -- between the treatment and Control groups must be regarded with caution, as they may have been a result of initial differences in ability to demonstrate the paper-pencil competencies. A note of caution must also be inserted regarding the findings involving paper-pencil exam performance. Test-retest reliabilities for the paper-pencil subscales ranged from -.30 to .90 (see Appendix B), suggesting considerable variability in the test's ability to measure

competencies consistently.

PA training effectiveness. The Physical Arrangement training proved to be highly effective. The PA treatment group made significant improvements in pre- to posttraining performance on six of the seven work sample and paper-pencil competencies. They also demonstrated posttraining performance significantly better than the combined control group (n=23) on six of the seven competencies. The only competency on which there was no improvement was work sample integration. The effectiveness of this module can be attributed to two primary sources. First, the module was rated by the six PA trainees who responded to the questionnaire as more than satisfactory in its ability to effectively train. It was well organized and presented concepts and training activities in a sequenced, logical order. Second, the PA trainer was rated by the six trainees as highly effective, able to ask thought provoking questions, helpful in aiding trainees to find new ways of applying what they learned, able to accept and value others' opinions, and well organized. In addition, he was instrumental in providing a comfortable learning environment for the trainees (meetings were held in the home of one of the trainees) and was rated as very enthusiastic and motivated. This enthusiasm and motivation seemed to be infectious, as evidenced by the PA trainee who opened her home to weekly training workshops, and another PA trainee who extended the skills taught in the module by incorporating the module's suggestion of arranging lofts in the classroom through building a movable platform, thereby adding an aspect of flexibility to the arrangement skills presented in the module.

LE training effectiveness. The Literature Experience training was

the least effective of the three OFT modules in its delivery of competencies. Although there were pre- to posttraining gains in performance on six of the ten LE competencies, only one of the gains achieved significance -- work sample diversity - BP. There were no significant posttraining differences between the LE and combined control groups.

The lack of positive statistical results for LE training were somewhat surprising. The Main Field Test of the module had indicated pre- to post-training gains on a paper-pencil exam significant at the .01 level (Rice, Taylor, and Chow, 1975). A good part of the contradictory OFT results may have been a function of the inexperience of the trainer, who had never served as a trainer before. The trainer was rated by a total of four LE trainees who responded to a feedback questionnaire as being less than satisfactory in training effectiveness. In addition, low ratings were given the trainer on training skills such as asking provocative questions, helping trainees discover ways to apply what they learned, accepting and valuing other's opinions, being well organized, and being enthusiastic and motivated. They also felt that the training workshops were dull and repetitive, and moved too slowly. Again, this may have been largely a function of the trainer. She reported that it had been difficult to understand parts of the module, and that the organization was not clear to her.

The module may also have contributed to the lack of improved performance. Its training effectiveness was rated very low by the LE trainees. It was criticized for its failure to present new material: "teachers already plan lessons around stories, so the module content was not new."

SP training effectiveness. The module Developing Children's Sense

Perception was the second most effective of the three OFT training modules. There were pre- to posttraining increases on eight of nine competencies, with improvement reaching significance on three paper-pencil competencies. SP trainee average performance was also better than the combined control group on eight of nine competencies, with significantly better performance on three paper-pencil competencies.

The module was rated by the three SP trainees who responded to a feedback questionnaire as a little less than satisfactory in its training effectiveness. Trainees seemed to feel that the module content was not diverse enough - that is, after they had completed the first few lessons and activities, the SP trainees felt that there was little to be learned from the module.

The SP trainer received low effectiveness ratings from SP trainees. She seemed to lack many of the same skills that the LE trainer lacked, e.g., the ability to ask thought-provoking questions, accept and value others' opinions, be well organized and display motivation and enthusiasm. There was some indication that the trainer did not have sufficient knowledge of the training module - there were trainee reports that she simply read from the module, and "often finished early and didn't know what to do..."

Summary of OFT Training Effectiveness

The results of the OFT research evaluation study have shown conclusively that the FLS training process can effectively deliver early childhood teaching skills. The module Physical Arrangement of the Classroom produced significant improvements in competency performance regardless of whether it was assessed via paper-pencil or classroom work sample measures. The findings of the study also suggest that there is a critical relationship

between the effectiveness of the module and its trainer. The Physical Arrangement trainer was consistently rated by trainees as more effective than the Sense Perception and Literature Experience trainers. This was in spite of the fact that all three trainers received the same amount of training by the module developers. Perhaps the greater experience of the PA trainer and his previous contact with FLS materials in a Far West Laboratory workshop accounted for his better performance as compared to the other two trainers.

It is difficult to partial out the effect of the trainer and training module in this study. The SP and LE were rated as less than satisfactory in their training effectiveness, and the LE training process was singularly unsuccessful in producing significant performance improvement, but was this a function of the module or the trainer? As originally conceived (Nimnicht and Johnson, 1973), the FLS modules were designed to be developed and installed by skilled Responsive Model trainers. Unfortunately, no provisions were made for the replacement of these trainer/developers as they left the FLS program. This is not to say that the FLS program cannot be successful without the experience and knowledge of a small core of persons present at its conception. Indeed, the Main Field Test of modules has repeatedly shown that the modules can produce changes without trainers extensively trained in its processes. Rather, it suggests that the philosophy and skills essential to FLS training be articulated and made accessible to trainers. The OFT results have begun to delineate what some of these skills might be, e.g., organization and questioning skills. In addition, the OFT findings suggest that trainers must possess certain characteristics such as enthusiasm, motivation, and willingness to accept

and value others' opinions.

In spite of the essential role that trainers play in FLS performance, it is also necessary to recognize the OFT findings regarding the weaknesses of the SP and LE modules. The SP module was fairly effective in its ability to produce significant positive changes in paper-pencil performance, but not work sample performance. This suggests that the SP module can deliver content knowledge regarding children's sense perception development, but not classroom teaching skills. Time permitting, it would be advisable to reassess the structure and scope of the module so that trainees would not find it repetitive and would be required to engage in more activities with children so that teaching skills are practiced and incorporated into a trainee's reportoire of teaching strategies.

The LE modules training effectiveness is somewhat difficult to assess because of the extreme confounding of trainer and module effects. The module was rated poorly by LE trainees, but this is at odds with the PFT and MFT findings. The LE trainer reported that she changed the order of module content presentation, not to make it more effective, but because she did not understand it. This may have generated confusion for the trainees. Nonetheless, the module itself would probably benefit from revision with redundant sections being modified or deleted.

Effects of Educational Level and Methods of Measurement

The findings from the comparisons of high and low education teachers suggests FLS training consistently improved the performance of low education teachers relative to high education teachers regardless of method-of-measurement.

Before training the high education group generated 13 appreciably

higher mean scores than the low education group ($t \geq 1.0$). These scores consisted of higher performance on five work sample skills (two of which reached significance) and eight paper-pencil skills (with five being significant). The low education group scored noticeably higher on only two paper-pencil competencies. These results suggest that low education teachers perform less well than high education teachers on teaching competency assessments measured by paper-pencil tests or classroom observations.

After receiving training there were 14 appreciable differences between low and high education group mean scores. This time, however, the low education group was appreciably higher on four work sample measures while the high education group was appreciably higher on three of the mean score comparisons (two of the latter achieved significance). On the paper-pencil measures the low education group was higher on three and the high education group on four (with one significantly higher).

This is considerably different from the pretraining results. The ratio of higher mean scores for high versus low education groups was 13:2 at pretraining and 7:7 at posttraining. The change in the ratio of significantly higher mean scores complements this trend; i.e., 7:0 at pretraining and 3:0 at post for the high and low education groups, respectively.

The differences between the two methods of measurement were negligible. At pretraining the high education group scored appreciably higher on five work sample and six paper-pencil measures relative to the low education group. At posttraining the low education group scored noticeably higher on one work sample measure than the high education group while the high education group scored appreciably higher on one paper-pencil measure than

the low education group. Thus, the pre- and posttraining performance on the two methods of measurement indicated that the large initial difference in favor of high education teachers did not exist after OFT training.

At this point the analysis turned to the educational differences within each treatment group. The low education teachers showed a total of eight appreciably higher mean scores; five were on work sample and four on paper-pencil measures with one from each method reaching significance. The high education teachers had a combined total of three appreciably higher mean scores than the low education teachers. Two of these were on work sample measures while the third was a significant difference on a paper-pencil measure.

Summary of OFT High and Low Education Teacher's Performance

Considered collectively, the pre- and posttraining results of teachers with relative high versus low amounts of formal postsecondary education suggests that initial differences between the groups do indeed exist. High education teachers tended to outperform low education teachers before training, but not after. When posttraining performance of high and low education teachers within particular treatments was examined, the low education teachers tended to consistently outperform high education teachers on both paper-pencil and work sample measures.

What could be accounting for these dramatic changes in performance? Perhaps the answer lies in the interaction between educational background and the FLS training process and content. Teachers with extensive education have been subjected to training processes and content in a methodical, systematic fashion. Thus, they have assimilated knowledge regarding theories of child development and teaching strategies directly

linked with these theories. FLS materials are accommodated in the context of this previous learning. Since many FLS competencies are based on developmental theory, the high education teacher's previous knowledge helps her/him display these competencies. This display is probably enhanced also by the test-taking skills obtained in the formal educational setting. However, the seeming advantages possessed by the high education teachers may constitute a double-edged sword in terms of teaching performance as a result of FLS training. The integration of FLS concepts and skills into the knowledge gained from previous educational training may produce conflicts or confusion in the learner's mind, resulting in a rejection or, at least, slower assimilation of knowledge or certain skills. In effect, prior training may make the learner less responsive to FLS training.

The teachers without such formal education, however, do not have the test-taking skills nor the theoretical underpinnings of their more educated colleagues. Consequently, they perform less well on paper-pencil exams and in classroom situations where theory-related competencies are demanded. On the other hand, the data accumulated throughout the FLS evaluation cycle indicated that many early childhood teachers with little postsecondary education tended to have considerable experience in the working with young children as parents or teachers. They have built their competencies upon an experiential foundation rather than a theoretical one. Thus, they can assimilate the concepts presented in the FLS materials without extensive accomodation of them with other theories. This lack of competition may result in increased ability to produce FLS theory-based competencies. This production may also be enhanced by the FLS training program. The process emphasizes learning in a reality-based context. Trainees practice,

modify, and expand their competencies through working with young children. This is highly similar to the trial-and-error basis that most low education trainees would presumably use to establish their teaching styles and strategies. Thus, they are readily able to grasp the skills delivered in FLS training and demonstrate them in a classroom situation. But what about the relative posttraining improvement of low education teachers on paper-pencil measures? Surely classroom-based training cannot substantially improve one's ability to perform on a written exam. Part of this improvement is probably due to pre- to posttest practice effects. In addition, however, the combination of FLS concepts and experiential activities presented in an easily read form may serve as a catylyst for the written expression of knowledge. It may be that teachers with relatively little education but extensive experience possess many of the concepts essential to the FLS but do not possess the language necessary to produce them in writing. The presentation of these concepts in what seems to be an easy-to-read, easy-to-understand format may solidify the low education learners conceptual base so that it is more easily expressed.

The posttraining improvement of low education relative to high education teachers is an encouraging indication of the FLS training program's effectiveness. One objective of the FLS is to serve as a career ladder for persons with the abilities but not the education to be certified as early childhood teachers. It appears that the FLS content, process, and presentation style may be facilitating accomplishment of this objective. On the other hand, the FLS is a prototype system with limitations, and it should be used as a tool that is complimentary or supplementary to other forms of training. Therefore, it is not in the least disconcerting that high education teachers do not seem to acquire FLS-related skills as

readily as low education teachers. High education teachers appear to undergo a process of "picking and choosing" competencies - that is, they seem to process FLS content and skills according to what they have previously learned and acquire only what they think to be valuable (indeed, trainee feedback summarized in Appendix C indicated that some high education teachers felt they were getting a "refresher" course). Thus, they do not perform relatively as well as low education teachers who seem to respond dramatically to training.

Recommendations for the FLS

The primary objective of the OFT was to determine whether FLS training modules could successfully deliver teaching skills to early childhood teachers. In essense, OFT findings regarding Physical Arrangement and Sense Perception training demonstrated that FLS modules can effectively alter teacher's performance on paper-pencil and classroom work sample methods-of-measurement. In addition, FLS training appears to eradicate many performance differences between teachers with limited or extensive education.

Although the OFT was not designed to evaluate specific strengths and weaknesses of the PA, LE, or SP modules, the findings did suggest that some refinement of the LE and SP modules might increase their effectiveness. Both modules need greater diversity in their content and activities. It is noteworthy that the work sample competencies of SP and LE personalization and SP diagnostics were not acquired by the trainees through training. Their lack of production is disappointing, for these competencies reflect the essence of the Responsive Education philosophy. They represent an ability to individualize learning experiences for the child, as they require

awareness of a child's presence in an activity (personalization) and ascertainment of his interests and behavior (diagnostics). The modules should perhaps emphasize skills such as these in their conceptual presentation of content and implement them as teaching skills through frequent inclusion in required classroom activities.

One of the most important outcomes of the OFT was its indentification of the critical link of the training module and the trainer. SP and LE training was not as productive as the PA training, and trainees using these modules consistently rated their trainers as less than satisfactory on several training skills. Thus, it is highly probable that a module's effectiveness is mediated by the trainer. One possible method of improving trainer effectiveness would consist of producing trainer manuals to accompany the FLS modules which require them. These manuals should emphasize particular training skills such as personalization of the training process, questioning behavior, and training preparation and organization. In addition, trainers should be familiar enough with the responsive philosophy of education to sustain a commitment to it. This might ensure the motivation and enthusiasm which seem to be necessary attributes of the effective trainer.

The trainer manuals should also include a segment devoted to evaluation and assessment of trainees. Differences in performance among teachers with high or low educational background suggest that high education teachers can be expected to outperform low education teachers before training but not after. Thus, training is important in the use of methods to evaluate trainees. In addition, an evaluator of trainees must be sensitive to individuals and diagnose their ability to perform in different modes

(i.e., written versus behavioral) and prescribe treatments for performance deficits. Of course, evaluation of this nature requires that a finely tuned assessment methodology be included as a part of the Flexible Learning System. This could involve a branched testing system whereby each trainee is administered sequential segments of a test and, based on her/his performance, is progressed through a module to a point where unacquired skills are presented. Assessment methods developed parallel to the segments within each module would be necessary for such an evaluation system to effectively operate. In addition, trainers would have to be equipped with the knowledge and skills necessary to constructively use the assessments as tools for trainee development of competencies.

The production of trainer manuals and the construction of a field-oriented evaluation methodology should ready the FLS for general dissemination. It is recognized that some of the training modules have difficulties inherent in them; however, discussion of these difficulties in the accompanying trainer manuals would contribute much to their alleviation by the trainer.

There is a considerable amount of research which, if conducted, could benefit the FLS as well as the field of early childhood education. One possible line of research would involve replication of the study presented in this report. Modifications in the design could improve the power of the study considerably. For example, the effect of trainers could be better accounted for if all trainers had approximately the same amount of prior training experience. In addition, the sample size could be increased to enhance the power and range of the statistical analyses. In the present study, a small sample was intentionally used to allow intensive examination and rigorous experimental control. The trade off,

however, was an inability to conduct analyses that may have lent more precision to the findings (e.g., multivariate analysis-of-variance). The competency measures also require improvement. Mean scores on a few of the work sample measures were extremely low. This could be due to non-production of the competencies, but it is also possible that the competency measures were not adequately assessing them. This is certainly the case with some of the paper-pencil measures. Their low reliabilities suggest a good deal of variability in what they were measuring over time.

Finally, the ethnographic approach to collection of supplementary data should be expanded and used consistently. A part of the OFT story lies untouched by the objective measures of competencies built into the research design. The measures, carefully deliberated and refined to the finest precision possible within the limits of our skills at research design, methodology, and analysis, are still only signs -- indications of teaching competence. Objective measures are like a few single threads of a rich, elaborate and intricate tapestry. To be sure, the threads were chosen with care. From all that we could tell by careful study of the modules, our instruments were focused at those points most likely to show changes in critical competencies. The odds seemed to be with us that if anything amenable to objective measurement showed changes, it would be the three or four key competencies identified by close reading and analysis of each of the manuals. However, the ethnographic data provided a tool for actually understanding the dynamic convolutions of teacher behavior within the study and how the study fit into the lives of those involved with it.

Additional research should be conducted to investigate differences

among trainers and how their performance might be improved. This could be accomplished by varying trainer use of the manuals and amount of trainer training.

As a general suggestion, it would be valuable to include an alternate treatment for a quasi-control group in research involving FLS materials and training. This would allow comparisons between the FLS and other models of teacher training.

The research proposed herein would benefit not only the FLS, but the field of early childhood education as well. It could begin to establish an empirically based hierarchy of validated teaching competencies. This would provide a firm basis for developing models of early childhood teaching which might truly benefit the growth of the young child.

References

Borg, W.R. The minicourse as a vehicle for changing teacher behavior: a three year follow-up. Journal of Educational Psychology, 1972, 63,6, 572-579.

Campbell, D.T., and Fiske, D.W. Convergent and discriminant validitation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 2, 81-105.

Cronbach, Lee J., Gleser, Goldine C., Nanda, Harinder, and Rajaratnam, Nageswari. The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles New York: John Wiley and Sons, Inc., 1972.

Dixon, W.J. (ed.) Biomedical Computer Programs, Berkeley: Univ. of Calif. Press, 1975.

Dunkin, Michael J., and Biddle, Bruce J. The Study of Teaching, New York: Holt, Rinehart, and Winston, Inc., 1974.

Gardner, Paul Leslie Scales and Statistics. Review of Educational Research, 1975, 45, 1, 43-57

Glass, Gene V., Willson, Victor L., Gottman John Mordechai. Design and Analysis of Time-Series Experiments, Univ. of Colorado, 1972.

Lord, Frederick M., and Novick, Melvin R. Statistical Theories of Mental Test Scores. Mass.: Addison-Wessley Publishing Co., 1974.

Medley, Donald M., Soar, Robert S., and Soar Ruth. Assessment and Research in Teacher Education: Focus on PBTE, Wash, D.C.: American Association of Colleges for Teacher Education, 1975.

Messick, Samuel The standard problem: meaning and values in measurement and evaluation. American Psychologist, 1975, 30, 10, 955-966.

Nimnicht, G.P., & Johnson, Jr., J.A. Beyond Compensatory Education, San Francisco: Far West Laboratory for Educational Research and Develpment, 1973.

Office of Child Development, DHEW Child Development Associate Training Grade, OCD 73-1065, Washington, D.C. 1973.

Taylor, T.D. & Chow, S. H. L. Main Field Test Report for Five Flexible Learning System Competency-Based Training Modules - Milestone B, Work Unit IV. Far West Laboratory for Educational Research and Development, 1974.

Ward, Beatrice A., Morine, Greta, and Berliner, David Assessing teacher competence in W. Robert Houston (ed.), Competency Assessment, Research, and Evaluation. A Report of a National Conference, Univ. of Houston, March 12-15, 1974.

APPENDIX A

PAPER PENCIL

AND

WORK SAMPLE

INSTRUMENTS, DEFINITIONS

AND PROCEDURES

PAPER-PENCIL

EXAM

100

A 2

Name: _____

Date: _____

NOTE TO THE LEARNER

This instrument has three sections with a total of 38 items. They deal with several aspects of teaching young children. We estimate that you will need about an hour to complete them; however, take as long as you need.

Thank you.

FLEXIBLE LEARNING SYSTEM

Far West Laboratory for Educational Research and Development

SECTION LE

Circle the best answer to questions 1-10.

1. A positive self-concept can be promoted in young children through:

   a. Selecting and reading books that have a calming influence on children.
   b. Selecting appropriate content and planning delivery of stories.
   c. Reading books whenever there is some extra time
   d. Plan reading as a group interaction activity.

2. When selecting stories to read to young children, it is important that the author depicts:

   a. Different people as basically the same.
   b. Heros and villains for children to identify with.
   c. Different people with self respect and dignity.
   d. Characters who represent the values of this society.

3. When new vocabulary words occur in a story:

   a. Definitions should be introduced before the story is read.
   b. Their meanings should be discovered by the children through context and illustrations.
   c. The teacher should change them to words already familiar to children.
   d. Definitions should be provided after the story is read.

4. When selecting stories to read to young children, it is important that:

   a. The story has sufficient dialogue.
   b. The story uses Standard English as a model.
   c. The language conveys the mood.
   d. Children can read the story themselves.

5. One of the first things to do in planning a literature experience is to:

   a. Choose ideas for the next literature experience.
   b. Decide which children to read to.
   c. Decide what lesson the children should learn from it.
   d. None of the above.

6. Which of the following is more important when reading a story to young children?:

   a. central location for story reading
   b. secluded location for story reading
   c. central and secluded location for story reading

7. The better children's books are likely to:

   a. Relate major events proceeding towards a climax.
   b. Have lively and colorful illustrations.
   c. Contain a substantial amount of dialogue and action.
   d. Depict unfamiliar experiences so children can learn from them.
   e. All of the above.

8. Story books which may have a positive impact on children are those conveying the feelings of:

   a. The individual's unique characteristics and contributions.
   b. The powerful forces of the natural order of things.
   c. The need to control emotions and feelings.
   d. The importance of the basic similarities of all people.

9. Which of the following is more important when reading stories to young children?:

   a. reading pace
   b. showing illustrations
   c. both a and b

10. When selecting stories to read to young children, it is important that the illustrations:

   a. do not compete with the text
   b. are simple and colorful
   c. make up at least one-third of the book
   d. occur on the same page as the text

11. From the list of activities below, select the three most important activities in planning a literature experience for young children. Number the activity you would do first as "1"; number the second activity as "2" and the third activity as "3":

   ___ Introduce definitions of unusual words.
   ___ Identify as many places in the story where discussions can occur.
   ___ Read the story to yourself.
   ___ Write preliminary ideas about the story.
   ___ Choose the most useful ideas for particular goals and individual children.

Read the story, The Fox and the Crow, and the instructions that follow, then answer questions 12-19.

The Fox and the Crow

A fox once saw a crow fly off with a piece of cheese in its beak and settle on a branch of a tree. "That's for me, as I am a fox," said Master Renard, and he walked up to the foot of the tree. "Good-day, Mistress Crow," he cried. "How well you are looking today: how glossy your feathers; how bright your eye. I feel sure your voice must surpass that of other birds, just as your figure does; let me hear but one song from you that I may greet you as the Queen of Birds." The crow lifted up her head and began to caw her best, but the moment she opened her mouth, the piece of cheese fell to the ground, only to be snapped up by Master Fox. "That will do," said he. "That was all I wanted. In exchange for your cheese I will give you a piece of advice for the future - do not trust flatterers."

Read the following four questions (12-15), that a teacher might ask her children about The Fox and the Crow, then decide whether each question was (a) asking the children about the theme, (b) determining if the children had background concepts or, (c) providing a purpose for listening to the story.

12. "Do you think that when people tell you nice things about yourself, you can always believe them?" This question is about: (circle one.)

    a. the theme of the story
    b. a background concept
    d. a purpose for listening

13. "What do you think of when I say 'bird'?" This question is about: (circle one.)

    a. the theme of the story
    b. a background concept
    c. a purpose for listening

14. "Is the bird in the story like the one you were thinking of?" This question is about: (circle one.)

    a. the theme of the story
    b. a background concept
    c. a purpose for listening

15. "What do you think the crow learned from what happened to her?" This question is about: (circle one.)

a. the theme of the story
b. a background concept
c. a purpose for listening

16. Which of the following teaching method(s) is more effective in providing The Fox and the Crow literature experience for young children? (circle one.)

a. "....but the moment she opened her mouth, the piece of _____ fell to the ground,..." (Teacher waits for children's response in _____.)

b. "A fox once saw a crow fly off with a piece of _____ in its beak ...." (Teacher waits for children's response in _____.)

c. both a and b are equally effective.

17. Which of the following teaching method(s) is more effective in providing The Fox and the Crow literature experience for young children? (circle one.)

a. "How did the fox promise to greet the crow if she would sing but one song?"
b. "How did the fox get the cheese from the crow?"
c. Both a and b are equally effective.

18. Which of the following teaching method(s) is more effective in providing The Fox and the Crow literature experience for young children? (circle one.)

a. "What would you have done in the crow's place?"
b. "What kinds of cheese did you think of?"
c. Both a and b are equally effective.

19. Which of the following teaching method(s) is more effective in providing The Fox and the Crow literature experience for young children? (circle one.)

a. "The crow lifted up her head and began to caw her best ..... What do you think will happen next?"
b. "Can you remember a time when someone tried to eat something you were trying to eat?"
c. Both a and b are equally effective.

SECTION PA

Instructions for questions 1-4: Read each question carefully, then circle the letter next to the best answer.

1. When "setting-up" a classroom, you should begin by:

   a. Deciding how many different learning areas the classroom should contain.
   b. Deciding what kind of activities there will be in the classroom.
   c. Deciding what kind of areas should be established for what kind of activities.

2. Mrs. Lawrence has five boxes of crayons for the children in her preschool room to use. The children always mix the crayons up when they are replacing them, even though Mrs. Lawrence has asked them to put the crayons back in the box they come out of. What is the most likely reason that Mrs. Lawrence has this problem?

   a. The children do not understand that they are to replace the crayons.
   b. The children do not like to follow rules.
   c. It is too difficult for the children to replace the crayons.

3. Which is the best way to provide for a variety of classroom activities for preschool children?

   a. Change the structure or purpose of the areas in the classroom occasionally.
   b. Eliminate some areas and add others occasionally.
   c. Change the location of areas occasionally.

4. A small group of children working with Ms. Potter need to get materials such as paste, scissors, and so on, that are stored across the room in an area where a few other children are working on individual art projects. What should Ms. Potter do to make sure the children in her activity can get the materials they need?

   a. She should divide the materials between the two areas: half in one area, half in the other.
   b. She should leave the materials in their regular place and let the children in her activity go get them.
   c. She should ask the children working on individual art projects to join in the activity she is going to carry out.

5. List four (4) physical characteristics that should be considered when arranging an area.

   a.
   b.
   c.
   d.

5. Match each statement on the right to an area described on the left.

___ Private area: A place for children to physically "drop-out" of the classroom.

___ Individual work area: A place where a child can work on his own projects.

___ Learning center <u>with</u> surfaces for sitting work.

___ Learning center <u>without</u> surfaces for sitting work.

___ General classroom activity area: A place where a number of different activities are scheduled.

___ Large-group free area: A place where activities involving the whole class are conducted.

A. Accommodates 3-6 children

B. A large open area in the classroom.

C. Accommodates 1-2 children and has no tables or chairs.

D. Children work standing up or sitting on the floor.

E. A small desk or table with one or two chairs.

F. Accommodates 5-12 children and is not set up for any particular activity.

7. Suppose that a teacher wants to allow 6-12 children in an area. List 3 ways of limiting traffic in the area without reducing the number of children the area can accommodate.

1. _____

2. _____

3. _____

8. Mr. Adams has picked up some carpet remnants for the tile floor in his classroom. He has just enough for two areas. The best combination of areas to use them in would have the activities of: (circle one.)

a. Science (planting and growing plants) and, reading (reading books, looking at pictures).
b. Cooking (measuring, pouring, cutting, mixing, etc.) and, dancing.
c. Dramatic Play (drama, puppet play, "dress-up," etc.).
d. Block Play (building, stacking, etc.) and, woodworking (cutting, sawing, pounding, etc.).

9. Mrs. Lee would like to have an area in her preschool classroom where each morning she could hold a 30 minute "cooking class." When she's not conducting a cooking class with the children, she wants the area to also serve as a "house play" area where children can play with the cooking utensils.

Answer the questions below by checking the statement that best describes how you would arrange this area. (Circle one.)

How large would you make the area?

a. Only big enough for one or two children.
b. Big enough for 3, but small enough so that no more than 6 could use the area.
c. Big enough for 5, but small enough so that no more than 12 could use the area.
d. Large enough so that most of the class could use the area.

What furnishing would you put in the area?

a. Counters that children can stand and work at.
b. Tables and chairs that children can sit and work at.
c. Counters and tables with a few chairs.
d. Any of the above.

How would you store materials (pots and pans, etc.) in the area?

a. The materials should be stored on open-shelves so that the children can see what they want and get it on their own.
b. The materials should be stored in high cupboards so that the teacher can get out only what she wants the children to have.
c. Some materials should be stored in high cupboards and some on open shelves.

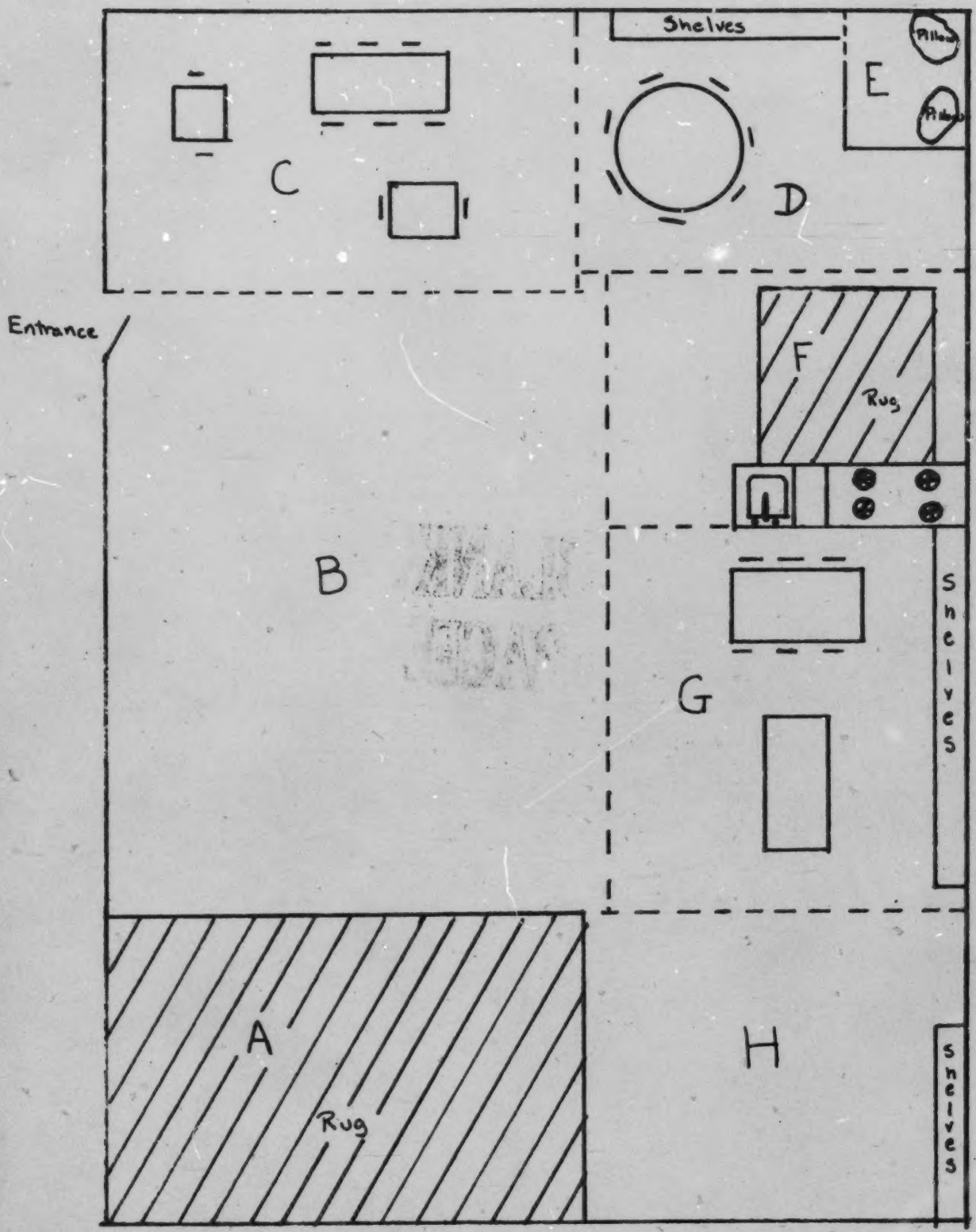
Should the area be open or closed?

a. The area should be open so that children can enter and leave.
b. There should be some barrier to separate the area from other classroom areas.
c. It doesn't matter if the area is open or closed as long as children know where it is.

10. The best way to describe areas in a classroom so that organization is indicated is to describe each area according to: (circle one.)

a. The activities that go on in it.
b. The type of materials that are used in it.
c. The characteristics of its arrangement.
d. How often the children use it.

Ms. Gill's Classroom



11. On the opposite page is a map of Ms. Gill's classroom. Each letter inside the dotted lines represents an area. Study the map and match the areas on the map with the activities that Ms. Gill wants to have.

Activities Ms. Gill wants:	Areas in classroom (put in the letter from the map of the classroom):
1. reading	_____
2. individual play or rest	_____
3. art	_____
4. tumbling and dancing	_____
5. individual projects	_____
6. block play	_____
7. housekeeping	_____
8. free-play	_____

SECTION SP

1. Suppose you had the following six, small, covered cans:

   Can No. 1, half-filled with pebbles.
   Can No. 2, half-filled with beads.
   Can No. 3, half-filled with sand.
   Can No. 4, half-filled with pennies.
   Can No. 5, half-filled with water.
   Can No. 6, empty.

   a. Describe three ways that you could have children group the cans.

   b. List two ways that you could have the child use the cans to make predictions about how they differ.

2. If you were observing a child to find out to what extent he uses his sense of touch, list at least three kinds of behaviors you would look for.

3. Look at the paper you are writing on, list all the words you can think of that describe it.

4. What are two important reasons for generating a wordlist for each of the senses (i.e., hearing, taste, etc.)?

5. List at least four criteria that a sensory exercise for children should meet.

6. The order of importance of the five senses is:

   a. touch, smell, taste, sight, hearing
   b. sight, touch, hearing, smell, taste
   c. sight, hearing, touch, smell and taste
   d. hearing, sight, smell, touch and taste
   e. none of the above

7. The best way to determine whether sensory development materials are appropriate for children is to: (circle one.)

   a. Make sure they match the activity they're intended for.
   b. Observe the children using the materials.
   c. Have a large variety of materials for each activity.
   d. Change the materials every two or three days.

8. Diagnosis of children's sensory development is most useful to: (circle one.)

   a. test children's performance
   b. assign grades
   c. determine the child's level of involvement
   d. determine sensory problems the children may have

9. Match one statement that a teacher could make during a sensory activity, with a cognitive skill that it would help develop in the child (the first one is done for you as an example):

Teacher Statements:		Children's cognitive skills:
a. "Which one is rough?"	____ 1.	Focus attention
b. "Put the rough ones here and the smooth ones here."	a 2.	Discriminate
c. "Which is more rough?"	____ 3.	Order
d. "Why did this happen?"	____ 4.	Sort
e. "Touch this."	____ 5.	Describe
f. "What should happen next?"	____ 6.	Compare
g. "What makes this sound?"	____ 7.	Name
h. "Put this one with one that is the same."	____ 8.	Match
i. "How does this feel?"	____ 9.	Explain Causality
j. "Put these objects in a line from hard to soft."	____ 10.	Predict

WORK SAMPLE

ON-SITE DATA COLLECTION PROCEDURES

1. Call teacher previous night to confirm visit and remind her to bring the learning experience materials.

2. Before departing for the data collection site, double check your equipment to ensure that everything necessary is accounted for.

   a. tape recorder with batteries and cord
   b. 3 audiotape cassettes
   c. extension cord
   d. clipboard
   e. quadrille paper for floor-plan diagram
   f. three sharp pencils and red/blue pencils
   g. classroom interview forms
   h. posttask interviews for literature and senses & perception experiences
   i. "incidental note" sheets
   j. camera
   k. 3 rolls film
   l. flash bulbs or cubes
   m. pencils and books
   n. new and extra books and materials
   o. stick-on labels

3. Arrive at the site a few minutes prior to the scheduled meeting time with the Research Study Teacher (the teacher will have received the learning experience materials one week before the first observation. In addition, she will have received an explanation of her role and the R.A.'s role on data collection days).

4. Introduce yourself and tell the teacher you are going to familiarize yourself with the classroom and test your equipment for about ten minutes.

5. Sketch 5-minute floorplan; Identify Area "A"

6. Label audio-tape cartridge and record identifying information. Listen to tape for quality.

7. Collect the classroom data using the Classroom Interview Instrument.

7a. Instruct teacher on operation of tape recorder and provide extra tape.

8. After completion of the Classroom Interview, ask the teacher where and when she would like to conduct the two learning experiences.

9. Before the children arrive:
   A. Label floorplan in blue letters
   B. Photograph each area
   C. Write film #'s in red letters
   D. Label film: RA#__/ Teacher #_____/ Observation #__/ Roll #___

10. Remain in the classroom for 10 minutes after the Classroom Interview is completed. Note any comments made by the teacher concerning the learning experiences or reference to your presence in the classroom (on "Notes").

11. Label the Classroom Interview tape (do not use the teacher's name, only her I.D. Number) and replay the tape to ensure quality and completeness. Check the floorplan diagram for accuracy.

12. Prepare your equipment for recording the first learning experience approximately 10-15 minutes before it is to begin. Label the new tape by learning experience (L.E. or S.P.), site, date, your I.D. #, and the teachers' I.D. #.

13. The recorder should be positioned on a flat, stable surface near to the teachers' seat. Use cord if possible. Test the equipment to ensure proper functioning.

14. Start the recorder as soon as the teacher and her students have situated themselves (remember, it's better to start taping too early rather than too late). Leave room.

15. After the first experience, check the quality and completeness of the tape. If there are problems, ask the teacher to re-do the experience with a different group of children.

16. As soon as possible, after 1st experience, conduct post task interview: Turn over tape; label tape (orally) with teacher #, RA #, Date, "L.E./S.P. Interview"; record interview, don't use teachers' name on the tape. Check the quality and completeness of the tape.

17. Prepare the equipment for the second learning experience by putting on a new tape, labeling it and making sure the recorder is still operating properly.

18. Check the quality and completeness of the taping as soon as possible after the learning experience. Turn over tape; label tape (orally) with teacher #, RA #, Date, "L.E./S.P. Interview": record interview, check for quality and completeness.

19. After the children leave, conduct the "Book Selection" and "Pencil Uses" interview. Check "Four Books Data Sheet" and tape of pencil uses for quality and completeness.

20. Before leaving the classroom, check that all equipment is accounted for, all tapes and diagrams are labeled, and generally that everything is completed and in order. Write sequence and timing of experiences and post task interviews in Notes.

21. Remind the teacher that the same process will occur one week from now. Give her the materials she needs to prepare the new learning experiences with. Take the old materials with you. Thank her for her cooperation.

POST TASK INTERVIEW

1. Names of children.

2. What do you think that each of the children got out of this experience?

3. Would you like to use this book/activity again? If no, why not? If yes, what would you do?

WORK SAMPLE

CLASSROOM INTERVIEW

Observation _____ Teacher # _____
Research Assistant #_____ Date_____

1. Ask teacher to take you on a guided tour of the classroom (where children do activities) proceeding in a clockwise fashion.

2. Tell teacher that you need to ask many questions and to save time, please try to answer them as briefly as possible.

3. Tell teacher that even though a question may seem irrelevant or repetitive, please try to answer it using best judgment or best guess.

For Each Area

1. What activities do children do here?

2. Do they do the same thing(s) anywhere else? If yes, where? If no, why not?

3. How many children can use this area at one time?

4. How many chairs are usually available?

5. Are the materials that the children use here stored here? Do the children get them themselves?

6. Is this area new or relocated? Has the furniture been rearranged since September? If yes, within the last 6 weeks?

AREA LETTER	ACTIVITY AREA	PHOTO (film number)
A		
B		
C		
D		
E		
F		
G		
H		
I		
J		

DATA CHECKLIST

Research Assistant: _____ Teacher Name: _____
Research Assistant #: _____ Teacher #: _____
Observation #: _____ Date: _____

1. Floorplan ____
2. Classroom Interview Tape ____
3. Classroom Interview Form ____
4. Photos ____
5. Learning Experience Tape ____
6. Learning Experience Interview Tape ____
7. Sense Perception Tape ____
8. Sense Perception Interview Tape ____
9. Notes ____
10. Book Form ____
11. Pencil Uses Tape ____
12. Pick up materials ____

FOUR BOOKS: DATA SHEET

Teacher #: _____ Date: _____
Observer #: _____ Time: _____

ORDER
(Best = 1, Next = 2, Next = 3, Worst

Krauss, Robert - "Whose Mouse Are You?" _____

McCloskey, Robert - "Make Way for Ducklings" _____

Merriam, Eve (Adapter) - "That Noodle-Head, Epaminoudas" _____

Rey, H. A. - "Curious George" _____

PHYSICAL ARRANGEMENT WORK SAMPLE

DEFINITIONS AND RATING PROCEDURES

PHYSICAL ARRANGEMENT WORK SAMPLE

RATING PROCEDURES

1. Review Floorplan.

2. Arrange photo sheets by alphabetical order.

3. Rewind tape to start position on side 1.

4. Complete procedures a through e for each area, beginning with Area A.

   a. Review photos for Area A in relation to Floorplan.

   b. Listen to interview on Area A only while looking at Floorplan and photos of Area A. Note responses on Rater's Note Sheet.

   c. Define type of area and write letter next to appropriate type on Balance Rating Sheet.

   d. Write area letters and # of areas on Stability Score Sheet.

   e. Listen again to Classroom Interview while looking at photos to score Quality of only those areas listed as types I, II, and III.

5. Use Rater's Note Sheet, Floorplan, and photos to make best judgments for Integration Score (you may refer to Classroom Interview sheet and tape if you wish).

PHYSICAL ARRANGEMENT CLASSROOM INTERVIEW RATING SHEET

Teacher # _____ Observation #__ Rater #___	Area:						
1. What activities do children do here?							
2. Do they do the same thing(s) anywhere else?							
If yes, where? If no, why not?							
3. How many children <u>can</u> use this area at one time?							
4. How many chairs are <u>usually</u> available?							
5. Where are the materials that the children use here?							
Do the children get them for themselves?							
6. Is this area new or relocated?							
Has the furniture been rearranged since September?							
Within the last 6 weeks?							
JUDGMENTS: a. Is this a NOISY area?							
b. Is this a QUIET area?							
c. Is this a MESSY area?							
d. Is this a FRAGILE area?							

PHYSICAL ARRANGEMENT

BALANCE RATING SHEET - DEFINITIONS

Learning Areas

Places in the classroom which house child initiated activities that are available for children to engage in everyday. It can be distinguished as with and without seating.

1. Learning Area Without Seats - learning center with either a purpose or specific activity area in which children engage in child initiated activities. Simultaneous activities may take place - such as puzzles, books, manipulative games without working surfaces for sitting work.
2. Learning Area With Seats - a learning center to serve specific activities such as manipulative math, reading, which has working surfaces such as tables and chairs. This area is arranged in such a way that children have free access to the materials they use there.
3. Private and Individual Areas - places in the classroom where one or two children can be by themselves. They can be distinguished as without or with seating.
4. Private Area - a small area where one or two children can go for privacy or isolation from the rest of the classroom. It has no working surfaces.
5. Individual Area - a desk or table where one or two children can go to work on something by themselves.

Group Areas

Places in the classroom at which groups of children can participate in a variety of activities during the week. They may include special activities. They can be distinguished as with or without seating.

6. Large-group Areas - a large open space with no specific function where between 12 children and the whole class could sit on the floor. Materials are brought to this area.
7. Free-group Work Area - No specific function but contains working surfaces with seating for 5-12 children. Materials are usually brought to this area.

PHYSICAL ARRANGEMENT

BALANCE RATING SHEET

Teacher #_____ Observation # _____
Rater # _____

Directions: The purpose of this rating sheet is to identify the primary type of use intended by the teacher in setting up each area of the classroom. Write the area letter for each type of area. After all areas have been tallied, use the criteria for that type of area to determine the score.

LEARNING AREAS are defined as places in the classroom which house child-initiated activities that are available for children to engage in everyday. Types I and II can be distinguished as without and with seating.

		Area Letters
I LEARNING AREA - WITHOUT SEATS	area without working surfaces for sitting work.	
II LEARNING AREA -	area with working surfaces for sitting work.	

PRIVATE AND INDIVIDUAL AREAS are defined as places in the classroom where one or two children can be by themselves. Types III and IV can be distinguished as without and with seating.

		Area Letters
III PRIVATE AREA	- a small area just big enough to afford one or two children privacy and isolation from the rest of the classroom. It has no specific function and does not contain working surfaces.	
IV INDIVIDUAL AREA	- a desk or table with one or two chairs where one or two children can go to work on something by themselves.	

GROUP AREAS are defned as places in the classroom at which groups of children can participate in a variety of activities during the week. These may include special projects which are not available everyday. Types III and IV can be distinguished as without and with seating.

		Area Letters
V LARGE GROUP AREAS	- a large open space with no specific function where between 12 children and all the class could sit on the floor.	
VI FREE GROUP WORK AREA	- an area with no specific function but contains working surfaces with enough seating to accommodate 5-12 children.	

PHYSICAL ARRANGEMENT

QUALITY RATING SHEET - DEFINITIONS

1. Single Purpose Area - an area in which only one activity can be going on at a time. Example (dramatic play).

2. Defined by Physical Boundaries - at least 2 sides of the area must be enclosed by room dividers, bookcase, screen, shelf, etc.

3. Visually Isolated - the physical boundaries must be at least four (4) feet tall, to enable adults but not children to see over. Usually it will be 3 sides, but can be two sides when the area is located in the corner of the room.

4. Isolated to Through Traffic - area is either located in a corner of a room, has a common entrance and exit, tables at angles, or area is away from main traffic patterns (main entrance and bathrooms).

5. Organized and Orderly Storage - most of the materials are grouped according to function, materials used in the same type of activity are next to each other, not separated.

6. Availability of Materials - the children can get the materials for themselves, and they are stored in the same area.

PHYSICAL ARRANGEMENT

QUALITY RATING SHEET

Teacher # _____ Observation # _____
Rater # _____ Score: _____

Directions: The purpose of this rating sheet is to determine how many quality criteria are met by areas of each type. Write the area letters for each type of area - as on the Balance Rating Sheet. Add the total for each area.

I LEARNING AREA WITHOUT SEATING AREA LETTERS: [ ][ ][ ][ ][ ]

1. This is a single purpose area. [ ][ ][ ][ ][ ]
2. This area is defined by physical boundaries. [ ][ ][ ][ ][ ]
3. This area is visually isolated. [ ][ ][ ][ ][ ]
4. This area is isolated to through traffic. [ ][ ][ ][ ][ ]
5. Three to six children can occupy this area. [ ][ ][ ][ ][ ]
6. No chairs are in this area. [ ][ ][ ][ ][ ]
7. Materials used in this area are accessible to the children. [ ][ ][ ][ ][ ]
8. Materials used in this area are visually evident to the children. [ ][ ][ ][ ][ ]
9. Materials used in this area are stored in an organized and orderly fashion. [ ][ ][ ][ ][ ]
10. Materials available in this area are supposed to be used here. [ ][ ][ ][ ][ ]

AREA TOTALS: [ ][ ][ ][ ][ ]

II LEARNING AREA WITH SEATING — AREA LETTERS: [ ][ ][ ][ ][ ]

1. This is a single purpose area. - - - - [ ][ ][ ][ ][ ]
2. This area is defined by physical boundaries. - - [ ][ ][ ][ ][ ]
3. This area is visually isolated. - - - - [ ][ ][ ][ ][ ]
4. This area is isolated to through traffic. - - [ ][ ][ ][ ][ ]
5. At least three and no more than six children can occupy this area. - - - - - - [ ][ ][ ][ ][ ]
6. There are three to six chairs in this area. - - [ ][ ][ ][ ][ ]
7. Materials used in this area are accessible to the children. - - - - - - [ ][ ][ ][ ][ ]
8. Materials used in this area are visually evident to the children. - - - - - - [ ][ ][ ][ ][ ]
9. Materials used in this area are stored in an organized and orderly fashion. - - - - - [ ][ ][ ][ ][ ]
10. Materials used in this area are stored here. - - [ ][ ][ ][ ][ ]

AREA TOTALS: [ ][ ][ ][ ][ ]

III PRIVATE AREA — AREA LETTERS: [ ][ ][ ][ ][ ]

1. This is a multipurpose area. [ ][ ][ ][ ][ ]
2. This area is defined by physical boundaries (such as room dividers, partitions, bookcases). - - - [ ][ ][ ][ ][ ]
3. This area is visually isolated. - - - - [ ][ ][ ][ ][ ]
4. This area is isolated to through traffic. - - - [ ][ ][ ][ ][ ]
5. Only one or two children can occupy this area. - - [ ][ ][ ][ ][ ]
6. No chairs are in this area. - - - - [ ][ ][ ][ ][ ]
7. No materials are available in this area. - - - [ ][ ][ ][ ][ ]
8. No materials are stored in this area. - - - [ ][ ][ ][ ][ ]

AREA TOTALS: [ ][ ][ ][ ][ ]

PHYSICAL ARRANGEMENT

INTEGRATION SCORE SHEET - DEFINITIONS

Noisy Activities Area - an area in which activities take place which would be considerably louder than reading, puzzles. Noisy activities would include block play, music and dance, dramatic play.

1. Is this noisy area away from, not adjacent to, a quiet area? Is this noisy area far enough away not right next to. Example: at the other side of the classroom.
2. Size restriction - an area limited to a certain number of children.
3. Quiet area - an area in which no loud noise takes place. Example: reading, a private area, individualized work area.
4. Messy activity areas - an area with activities such as; art, cooking, gluing, pasting, water play.
5. Messy area separated from a fragile area - it can be separated by a physical boundary. A fragile area - materials in an area that have the potential of being damaged by something messy.
6. Is this messy area adjacent to an available water source? If the water source is located in close proximity to messy area. There may be more than one messy area close to a water source - Example: art, cooking, painting.
7. Appropriate flooring -
   a. Noisy area - carpet, area rug
   b. Messy area - tiled, plastic cloth under table

PHYSICAL ARRANGEMENT

INTEGRATION SCORE SHEET

Teacher #_____ Observation #_____
Rater #_____ Score: _____

Letters of noisy activity areas:							
1. Is this noisy area separate from, not adjacent to, a quiet area?							
2. Is this noisy area on a carpeted or soft floor surface?							
3. Does this noisy area have size restriction?							
Totals for each area							

Letters of messy activity areas:							
1. Is this messy area separated from a fragile area?							
2. Is this messy area adjacent to an available water source?							
3. Is this messy area on appropriate or protected flooring?							
Totals for each area							

PHYSICAL ARRANGEMENT

SUMMARY SCORE SHEET

Teacher #_____ Observation #_____

Rater# _____ Date: _____

BALANCE SCORE:

# of Area Letters	Maximum # of Points	Score
I _____	2	
II _____	2	
III _____	1	
IV _____	2	
V _____	1	
VI _____	1	

Total = ☐ Balance Score

QUALITY SCORE:

	Raw Score (X 8) =	Type Score	# of Areas of each type
I			
II	(X 8) =		
III	(X 10) =		
Total	(Sum of Type Scores)	÷	(Sum of Number of Areas)

= ☐ Quality Score

STABILITY SCORES:

Total since September _____

Total within last 6 weeks _____

= ☐ Stability Score

INTEGRATION SCORES

	NOISY AREAS	MESSY AREAS
A. Highest scored area.		
B. The additional number of areas that met all three criteria.		
C. The sum of A and B.		

= ☐ Integration Score

Total equals C. for noisy plus C. for messy

LITERATURE EXPERIENCES WORK SAMPLE

DEFINITIONS AND RATING PROCEDURES

LITERATURE EXPERIENCES

RATING PROCEDURES

1. Listen to audio-tape of Story-Reading Experience (LE) while following along in an appropriate book. Tally teacher behavior on Diversity Rating sheet.

2. Listen to audio-tape of Post-task Interview. Write in children's names on Personalization Rating sheet. Tally teacher's description of levels of skill and involvement of each child.

3. Book Selection procedures are in the Sense Perception section with the Pencil Uses Interview Procedures.

LITERATURE EXPERIENCE

DIVERSITY RATING - DEFINITIONS

BEFORE - Includes all discussion up to first page of story.

1. Verbal statement made by teacher regarding theme or background concepts.

   Theme - message author is trying to convey.

   Ex. Birthday Presents (BP) the most important and lasting gifts are intangible.

   Ex. Snowy Day (SD) what a young child can do in the snow.

   Background Concepts - deal with initial climate of story; such as setting, tone or first event of story. It should relate to theme.

   Ex. BP: This book is about birthday presents.

   Ex. SD: This book is about a little boy playing in the snow.

   This may include reference to a relevant class activity.

   This may not include statements regarding vocabulary, story, climax, or final events of story.

2. Questions initiated by teacher regarding theme or background concepts.

   Theme Ex. BP: What presents did you get for your last birthday?

   Ex. SD: What is the little boy doing on the cover?

   Background Concepts (see #1)

   Ex. BP: Who had a birthday this month?

   Ex. SD: Did you ever see snow?

   These questions may not include premature or minor issues and concepts pertaining to plot.

   Ex. BP: What is a song?

   Ex. SD: What happens to snow when it melts? (premature).

3. Teacher responds to child's participation with positive acceptance.

The following responses are coded:

1. Teacher credits all child's credible ideas or words even if they are not exactly the same as author's version.

2. Teacher acknowledges the extra explanations, alternatives, or questions that child might offer during story.

Teacher can respond with:

a. Really, that's nice.
b. Repetition of child's words.

Regarding Teacher's response to child's questions:

a. Teacher should answer questions quickly if it seems like method of inductive reasoning would spoil flow of story.

b. Teacher says let's see what happens if we read on.

c. If it doesn't interrupt story flow, teacher can guide child by asking another question.

4. Teacher states a specific purpose for listening. Something child hopes to find out or decide upon as a result of listening.

Ex. BP: See if anything about this boy's birthday is like yours.

Ex. SD: Let's see why this book is called the Snowy Day.

Questions that would stimulate interest immediately might relate to the title, the cover (Ex. SD: Let's see what the boy on the cover does in the snow), the title page.

Inappropriate questions:

a. Insult listener by focusing on detailed information.

b. Provide story content prematurely. (Ex. SD: What will happen to snow ball?)

DURING - Teacher starts with first page and reads up to last page of story.

1. Teacher gives children the opportunity to make predictions.

   Predictions - Child is asked to hypothesize upcoming events and consequences.

   What will happen next?
   How will it end?
   How will it go?

   Ex. SD: (p. 15) What will happen when Peter hits the tree?

   Ex. BP: What happened to the truck? The child must have enough information to make a logical prediction.

2. Teacher gives opportunity to supply closure.

   Closure - The listener is asked to anticipate exact words or ideas to follow on the basis of information presented, thus "closing the sentence."

   The teacher reads right up to the missing link and pauses for children's response. This pause must come at a point where there is enough information available for a reasonable prediction.

   Opportunities for closure are:

   1. Rhyming words or alliteration.
   2. Recurring words, phrases or sentence or group of senteces.
   3. Anticipation of word on basis of cause and effect relationships.
   4. Anticipation of grammatical forms such as tense, number or gender.

3. Teacher gives children opportunity to offer alternatives or explanations to problems and situations posed in story based on their own logic and rationale.

   Ex. How else might story have ended?
   What else could they have done?

   Ex. SD: Why do you think he felt sad?

   Ex. BP: Why did grandfather choose to give a song?

   Inappropriate questions:

   1. Deal with trivial information - Ex. SD: How many buildings did Peter see from his window?

2. Information that will be readily available in the text.

3. That interfere with thoughtful appreciation of the whole (i.e., climactic point).

AFTER - All discussion that takes place after reading of story.

1. Teacher guides children to make application to their own lives by recalling past experiences.

   A. Discussion - Ex. BP: What about Joseph's birthday; is it like yours?

   Ex. SD: Have you ever played in the snow?

   B. Class projects (i.e., art or role-playing).

2. Teacher guides children to make application to their own lives by projecting themselves into new experiences. This may include reference to a projected class project.

   A. Discussion - Ex. BP: Think of someone you live with or a special friend. What special gifts like grandfather's song, could you give that person?

   B. Class projects (i.e., art or role-playing).

3. Teacher responds to child's participation with positive acceptance (see BEFORE #3).

4. Teacher makes a verbal statement regarding fulfillment of purpose for listening.

LITERATURE EXPERIENCE

DIVERSITY RATING

Teacher #_____ Date: _____
Observation #_____ Score: _____
Rater #_____

Directions: Please tally teacher behaviors before, during, and after story reading experience.

Score: Number of categories with entries. ☐

BEFORE:

1. Teacher makes a verbal statement regarding background concepts or theme of story. This may include reference to a relevant class activity. This may not include statements re vocabulary, story climax, or final events of story. ☐
2. Teacher asks children a question regarding background concepts or theme of story. ☐
3. Teacher responds to child's participation with positive acceptance. ☐
4. Teacher states a specific purpose for listening. ☐

DURING:

1. Teacher gives children the opportunity to make predictions. ☐
2. Teacher gives children the opportunity to supply closure. ☐
3. Teacher gives children the opportunity to offer alternatives or explanations. ☐
4. Teacher responds to child's participation with positive acceptance. ☐

AFTER:

1. Teacher guides children to make application to their own lives by recalling past experiences. ☐
2. Teacher guides children to make application to their own lives by projecting themselves into new experiences. This may include reference to a projected class activity. ☐
3. Teacher responds to child's participation with positive acceptance. ☐
4. Teacher makes a verbal statement regarding fulfillment of purpose for listening. ☐

LITERATURE EXPERIENCE

PERSONALIZATION RATING

Teacher:_____ Date:_____ Time:_____

Recorder:_____ Rater:_____

	CHILD'S NAME				
TOTAL SCORE _____	1	2	3	4	5
1. STORY: Teacher describes reaction of individual child to the story (e.g., "Tom knew what was going to happen next.") SCORE _____					
2. PERSONAL: Teacher describes reaction of individual child to personal experience (e.g., "Maria has been talking about her birthday all week.") SCORE _____					

SCORE = The total number of individual children described by the teacher in the Story and Personal sections above

TOTAL SCORE _____

SENSE PERCEPTION WORK SAMPLE

DEFINITIONS AND RATING PROCEDURES

SENSE PERCEPTION

RATING PROCEDURES

1. Listen to audio-tape of Pencil Uses Interview (PU), and tally sensory uses on Sensory Awareness Score Sheet.

2. Listen to audio-tape of Sense Perception Experience (SP), and tally matrix: teacher behavior x sense mode.

3. Listen to audio-tape of Post-Task Interview, and a) write in children's names, b) tally teacher descriptions of levels of skill and involvement of each child.

SENSE PERCEPTION

DIVERSITY RATING - CODING AGREEMENTS

Two overriding principles of coding are:

1. Comments by the teacher which invite the children to participate in some way in a sensory experience should be coded.
2. Questions which focus on the attributes of objects, as opposed to simply naming the objects, are to be coded.

Example: If the teacher asks a child to feel a bead, that is coded as "Make Statement - Touch." If the teacher then asks the child to guess what the object is, that question is not coded. The question would be coded if the teacher said, "What does it feel like?" Names of spices are considered attributes when used with "like", as "like cinnamon". If the teacher does not refer to a specific sense or discriminatory skill, do not code the comment.

When the teacher repeats a child's statement on some attribute, code statements in "Make Statement" and questions in "Ask Label."

Example: "Yes, it's brown." "Make Statement - Sight."
"It's hard?" "Ask Label - Touch."

Code each repetition by teacher, even if they appear in rapid succession, as "touch it - go on, touch it - did you touch it yet?" Each is another focus or invitation to participate, and would be coded as "Make Statement."

- If teacher uses additional materials, disregard all references to them.
- When in doubt as to how to code an ambiguous statement, code in more conservative block, i.e., "Ask Label" is more conservative choice than "Ask Pattern."
- To distinguish between "Ask Label" and "Ask Relationship"

  If the teacher asks a child to identify by opposites, i.e., "Is it hard or soft?" Code as "Ask Label - Touch."

  If the teacher asks for a comparison within a range of one attribute, i.e., "Which is harder?" Code as "Ask Relationship - Touch."

A question may not be coded in more than one place. If the teacher says, "What shape is it?" without specifying the sense to be used, code under "Ask Label - Sight." Between sight and touch, choose sight if not otherwise specified. If however, the question clearly refers to two sense modes, code under both of them, i.e., "Look at it and feel it and tell me what shape it is."

A question may be inferred by context to be in the sense mode that is being addressed in an overall activity.

Each teacher statement is to be coded, but five hatch marks per coding box is considered sufficient to indicate frequency of teacher behavior.

SENSE PERCEPTION

DIVERSITY RATING - DEFINITIONS

1. Make Statement - Touch

   Teacher focuses children on aspects of shape, size, weight, or texture.

   Example: "Feel how coarse this is."
   "It's very light, isn't it?"
   "It feels like sand."

2. Ask Label - Touch

   Teacher asks children to identify some attribute of beads or spices by feeling them.

   Example: "How does it feel?"
   "Does it feel soft?"

3. Ask Relationship - Touch

   Teacher asks children to compare or match attributes of beads or spices, as well as to discriminate by attributes, using the sense of touch.

   Example: "Close your eyes and tell me if these feel the same."
   "Does this one feel like that one?"
   "Pick out the one that's rough by using your fingers."

4. Ask Pattern - Touch

Teacher asks children to group, classify, sort or order beads or spices on ths basis of attributes discernable by touch.

Example: "Put all the smooth ones together."

"Copy my pattern of hard, soft, hard, soft."

"Pick out the ones that are green."

5. Ask Science - Touch

Teacher asks children to explain or predict some attribute of beads or spices by touch.

Example: "Why are these both so smooth?"

"How did it get so hard?"

6. Make Statement - Hearing

Teacher focuses children on the sound of the beads or spices.

Example: "Listen to this."

"Shake it and listen."

"Those beads are noisy."

7. Ask Label - Hearing

Teacher asks children to describe the sound of the beads or spices.

Example: "How does that sound?"

"Is it quiet or loud?"

8. Ask Relationship - Hearing

Teacher asks children to compare, match, or discriminate sounds.

Example: "Does this one sound like that one?"

"Which is the loudest?"

"Which one sounds like this?"

9. Ask Pattern - Hearing

Teacher asks children to group, classify, sort or order on the basis of sounds.

Example: "Make a group of the noisy ones."

"Can you tap out this rhythm?"

"Order these from loudest to softest."

10. Ask Science - Hearing

Teacher asks children to explain or predict the sound of spices or beads.

Example: "Why is it so quiet?"

"How does it make that sound?"

11. Make Statement - Sight

Teacher focuses children on the visual attributes of an object; invites participation in looking.

Example: "Look at that."

"This is a round one."

12. Ask Label - Sight

Teacher asks child to identify or name visual attributes of objects.

Example: "Is it red?"

"How does it look?"

"What color is it?"

13. Ask Relationship - Sight

Teacher asks child to compare/contrast, discriminate or match attributes or objects based on visual discrimination.

Example: "Which one is red?"

"Is there another round one?"

"Which one is darker?"

14. Ask Pattern - Sight

Teacher asks child to sort, order, group or classify objects, to create or perceive a pattern by sight.

Example: (in establishing a pattern) "Which one goes next?"

"Put these in order from big to little."

"Put the ones that look the same together - why do they look the same?"

15. Ask Science - Sight

Teacher asks child to explain or predict the properties of an object by using sight.

Example: "How would it look if we stepped on it?"

"Why is it red?"

16. Make Statement - Smell and Taste

Teacher focuses children on smell and taste attributes; invites participation by use of smell and taste.

Example: "Take a taste."

"It smells like flowers."

17. Ask Label - Smell and Taste

Teacher asks children to determine an attribute by using sense of smell or taste . Names of spices can be considered attribute labels.

Example: "Does it smell good?"

"How does it taste?"

18. Ask Relationship - Smell and Taste

Teacher asks child to compare/contrast, discriminate or match one object from others in the array using the senses of smell and touch.

Example: "Which one smells like flowers?"

"Can you find another one which tastes the same as this one?"

"Which one tastes better?"

19. Ask Pattern - Smell and Taste

Teacher asks child to group, classify, sort or order objects - to perceive or create a pattern using smell or taste.

Example: "Put all the ones that taste sweet together."

"Put all the ones that smell the same together - how do they smell?"

20. Ask Science - Smell and Taste

Teacher asks child to explain or predict the properties of an object by using the senses of smell and touch.

Example: "Will these beads smell different?"

"Which of these would taste like they smell?"

SENSE PERCEPTION EXPERIENCE

DIVERSITY RATING

Teacher #_____
Observation# _____ Date: _____
Rater: _____ Materials: _____

Directions: Tally Teacher behaviors which foster sense perception skills.

	TOUCH	HEARING	SIGHT	SMELL & TASTE	
1. Make Statement. Teacher focusses attention by making a statement or command (e.g., "This is a red one;" "Touch this.").	(1)	(6)	(11)	(16)	
2. Ask Label; identify, name, describe: Teacher asks child for verbal description or attribute label of single object (e.g., "How does it feel?" "What color is it?").	(2)	(7)	(12)	(17)	
3. Ask Relationship; compare/contrast, discriminate, match: Teacher asks child to evaluate relationship between one object and others in the array (e.g., "How is this one different?" "Which one is red?" "Which is more rough?" "Put this one with one that is the same.").	(3)	(8)	(13)	(18)	
4. Ask Pattern; group, classify, sort, order: Teacher asks child to arrange several objects in a pattern (e.g., "Put the rough ones here, and put the smooth ones here," "Put the ones together that are the same; why do you think they are the same?" "Place these objects in a line in order from hard to soft.").	(4)	(9)	(14)	(19)	
5. Ask Science; explain, predict: Teacher asks child for a scientific rationale (e.g., "How would it look if we stepped on it?" "Why is it red?"	(5)	(10)	(15)	(20)	

SENSORY EXPERIENCE

DIAGNOSTIC RATING

Teacher: Date: Time:
Recorder: Rater:

	CHILD'S NAME				
	1	2	3	4	5
1. SKILL: Teacher describes level of skill of individual child (e.g., "Bobby had a hard time distinguishing the brown from the wooden beads.") SCORE ___					
2. INVOLVEMENT: Teacher describes level of involvement of individual child (e.g., "Luisa really enjoys different textures.") SCORE ___					

SCORE = The total number of individual children described by the teacher in the Skill and Involvement sections above.

TOTAL SCORE _____

BOOK SELECTION AND PENCIL USES INTERVIEW PROCEDURES
(conducted during paper and pencil assessment)

1. Find a private interview area. Set up array of four books:

   a. Krouss, Robert - "Whose Mouse Are You?"
   b. McCloskey, Robert - "Make Way for Ducklings"
   c. Merriam, Eve (adapter) - "That Noodle-Head, Epaminoudas"
   d. Rey, H. A. - "Curious George"

2. Invite teacher to interview area. Write identifying information on Four Books: Data Sheet.

3. Ask teacher to look at the books (about 5 minutes per book) so as to arrange them in order from best to worst in terms of selecting them to read to the children in her class.

4. While teacher is reviewing books, set up audio-tape recorder. Record identifying information. Listen to tape.

5. Write book order on Four Books: Data Sheet.

6. Give teacher pencils. Ask teacher to:

   "Please think for a few minutes of the many possible uses of these pencils with young children." (pause)

7. Start recording: "Please tell me all the ways you can think of using these pencils with young children in your class." (Probe ONLY if no response): "What activities could you set up for children with these pencils?

8. When list is finished: "Can you think of any other possibilities?

9. Say "thank you."

SENSE PERCEPTION

AWARENESS SCORE

Sensory Awareness Score: _____

Touch: _____

Sight: _____

Hearing: _____

Smell: _____

Taste: _____

Non-Sensory Experiences: _____

Score:

100 X #Sensory Experiences listed

÷ Total #Experiences listed.

LITERATURE EXPERIENCE

SELECTION SCORE

Scoring procedure:

1. Rank order correlation between the subject's and the module developer's rating of the four books was calculated.
2. There are eleven possible correlations. They increase from -1.0 to 1.0 in even steps of 0.2.
3. Each subject was assigned a score from zero to ten.

WORK SAMPLE DATA COLLECTION PROCEDURES

On-Site Data Collection Procedures

One week before the first classroom data collection visit, each teacher in the sample was sent a packet of information and materials. The packet included a schedule for and procedures during the visits, as well as the book and sensory stimulus materials. Each classroom visit followed a standard sequence of On-Site Data Collection Procedures. These procedures began with a phone call by the visiting research assistant to the teacher the night before each visit to confirm the visit and remind the teacher to bring the materials sent the previous week.

In the morning, after the teacher and research assistant had met, the first task of the research assistant was to obtain information about the arrangement of the classroom. First, she/he sketched a floorplan of the classroom and invited the teacher to take her/him on a "guided tour" of the classroom. During the tour, the research assistant made an audiotape recording of the teacher's responses to the classroom interview (six questions regarding each activity in the classroom). After the interview was completed, the research assistant labelled the floorplan and took photographs of each area that was identified by the teacher. The second and third tasks of the research assistant were to obtain audiotape recordings of the teacher (a) conducting a story reading

learning experience and sense perception learning experience (each with a different small group of children) using the appropriate stimulus materials, and (b) responding to a post-task interview after each experience regarding children's participation. On the second and fourth occasions, the visitor conducted an additional interview in which the teacher was asked to rank his/her selection of a randomly ordered set of four books and to generate a list of ways in which a standard set of pencils could be used in activities with children.

After each data collection visit, the research assistant brought all work sample data back to the Laboratory for rating and scoring. The work samples and the measures obtained from them are listed in Table B.

Training for Data Collectors

Each of the data collectors had some previous classroom teaching experience, and some were working part-time concurrently as teachers in local children's centers. Their training was conducted over a one-month period in weekly three day group sessions at the Laboratory and in local preschools. As the research design involved a small sample at a considerable distance from the Laboratory, it was considered essential that all data be complete, accurate and of good quality. Thus, an intensive one-month training program was conducted at the Laboratory wherein each of the necessary procedures were introduced, tried out in the Laboratory, and tried out at a local field site.

The first week of training began with procedures for collecting information about the physical arrangement of the classroom. The research

assistants practiced drawing floorplans until they could produce one accurately within a five-minute time limit. They practiced taking photographs under the instruction of a professional photographer until they could provide a distance shot showing the relationship of any area to adjacent areas, an aerial shot showing an area's furniture arrangement, and a "close-up" showing its storage facilities and arrangement. They practiced interviewing one another until they had learned the basics of audiotape recording, had memorized the interview questions, and could conduct a responsive, yet structured, interview.

The second week of training included procedures for collecting audiotape recordings of teachers reading stories or conducting sense perception experiences with young children in classroom settings. Research assistants formed pairs to practice recording procedures. They were instructed on operating and trouble shooting audiotape recorders until they were equipped to handle most problems which might occur in the field. Training also included procedures for post-task, book selection, and pencil uses interviews.

During the third week the research assistant worked on integrating each of the procedures previously learned, and on learning procedures for organizing materials, equipment, and data.

The first meeting of the fourth week was devoted to an in-house rehearsal of the entire set of procedures, and the second meeting involved a dress rehearsal at a local pre-school.

Throughout the entire training process, regular opportunities were provided for group discussion and for individual and group feedback. For example, research assistants shared their floorplans and discussed

one another's tape recorded teacher interviews. In addition, time was spent discussing and role-playing the interpersonal relations likely to develop in communication between the sample teachers and the research assistants. This was done to maximize the likelihood of obtaining cooperation from the sample teachers.

Rating and Scoring Procedures for Work Sample Data

Physical Arrangement of the Classroom

The physical arrangement work sample data from each classroom visit consisted of a floor plan of the entire classroom, photographs of each area, and an audiotape recording of the teacher's responses to the Classroom Interview. The instruments used for rating the data produced a score for each of the three basic competencies identified for the module: Balance, Quality, and Integration.

Balance: The purpose of this measure was to determine whether teachers arranged their classrooms in accord with the proportion of six types of classroom areas recommended in the module. All the data available for each classroom area identified by the teacher were used as the basis for identifying the primary type of use intended by the teacher in setting up the area. The six primary types presented in the module were: learning area without seats, learning area with seats, private area, individual area, large-group area, and free-group work area.

A score was assigned to each area type according to whether it had the minimum number of area types to match the recommended proportions:

Area Type	Minimum Number
I. Learning Area Without Seats	2
II. Learning Area With Seats	2
III. Private Area	1
IV. Individual Area	2
V. Large Group Area	1
VI. Free Group Work Area	1

Quality: The purpose of this measure was to determine whether teachers arranged learning areas without seats, learning areas with seats, and private areas according to criteria presented in the module.

Learning areas without seating were given one point for meeting each of ten criteria:

1. This is a single purpose area.
2. This area is defined by physical boundaries.
3. This area is visually isolated.
4. This area is located to through traffic.
5. Three to six children can occupy this area.
6. No chairs are used in this area.
7. Materials used in this area are accessible to the children.
8. Materials used in this area are visually evident to the children.
9. Materials used in this area are stored in an organized and orderly fashion.
10. Materials available in this area are supposed to be used here.

Learning areas with seating were assigned one point for meeting each of ten criteria:

1. This is a single purpose area.
2. This area is defined by physical boundaries.
3. This area is visually isolated.
4. This area is isolated to through traffic.
5. At least three and no more than six children can occupy this area.
6. There are three to six chairs in this area.
7. Materials used in this area are accessible to the children.
8. Materials used in this area are visually evident to the children.
9. Materials used in this area are stored in an organized and orderly fashion.
10. Materials used in this area are stored here.

Private areas were given one point for meeting each of eight criteria:

1. This is a multipurpose area.
2. This area is defined by physical boundaries (such as room dividers, partitions, bookcases).
3. This area is visually isolated.
4. This area is isolated to through traffic.
5. Only one or two children can occupy this area.
6. No chairs are in this area.
7. No materials are available in this area.
8. No materials are stored in this area.

The ratings were summed for each of the three area types and multiplied by a factor of either eight or ten so that each area would have the same maximum possible score regardless of number of criteria. This was done so that quantitative differences in eligible areas due to differences in sizes of classrooms would not be reflected in the Quality score. An average was computed for each of the three area types. The sum of these averages was used as the Qualtiy score.

Integration: This measure indicated whether those areas judged as either noisy or "messy" met three criteria for being integrated into the classroom as a whole. The criteria for noisy areas were:

1. Is this noisy area separate from, not adjacent to, a quite area?
2. Is this noisy area on a carpeted or soft floor surface?
3. Does this noisy area have size restriction?

The criteria for messy areas were:

1. Is this messy area separated from a fragil area?
2. Is this messy area adjacent to an available water source?
3. Is this messy area on appropriate or protected flooring?

The integration score was the sum of criteria met for the best case in each of the categories of noisy and messy areas. An additional point was given for each area tied for best case; i.e. that met all three criteria.

Enriching Literature Experiences of Young Children

The work sample data for the Literature Experience Module competencies were collected on each classroom visit. They included an audiotape recording of the teacher reading a children's book (either Snowy Day or Birthday Presents) to a small group of children in the classroom, and an audiotape recording of the teacher's response to an interview conducted after the

story-reading experience. In addition, the Book Selection Data Sheet was completed on the scond and fourth visits, wherein the research assistant indicated the teacher's verbally expressed preferred order of four children's books. The instruments used for rating these data produced a score for each of the three basic competencies of Selection, Diversity, and Personalization which had been identified for the module. The procedures for deriving these scores are discussed below.

Selection: The Selection score was calculated from the Book Selection Data Sheet as the rank order correlation between the subject's and the module developer's rating of the four books. There are eleven possible correlations. They increase in even steps of 0.2. scores were assigned from zero to ten.

Diversity: The purpose of this measure was to determine how many of the twelve categories of criterion teacher behavior were performed and recorded during the course of the story-reading experience. Tallies were made for four categories of teacher behavior before a teacher began reading the story:

1. Teacher makes a verbal statement regarding background concepts or theme of story. This may include reference to a relevant class activity. This may not include statements re vocabulary, story climax, or final events of story.
2. Teacher asks children a question regarding background concepts or theme of story.
3. Teacher responds to child's participation with positive acceptance.
4. Teacher states a specific purpose for listening.

Tallies were made for four categories of teacher behavior during story-reading:

1. Teacher gives children the opportunity to make predictions.
2. Teacher gives children the opportunity to supply closure.
3. Teacher gives children the opportunity to offer alternatives or explanations.
4. Teacher responds to child's participation with positive acceptance.

Tallies were made for four categories of teacher behavior after the last page of the story has been read:

1. Teacher guides children to make application to their own lives by recalling past experiences.
2. Teacher guides children to make application to their own lives by projecting themselves into new experiences. This may include reference to a projected class activity.
3. Teacher responds to child's participation with positive acceptance.
4. Teacher makes a verbal statement regarding fulfillment of purpose for listening.

The Diversity score was computed as number of categories with entries. Thus, the maximum possible score was 12.

Personalization: The purpose of this measure was to indicate the extent to which a teacher produced a description of the personal reactions of individual children to the story-reading experience. The source of data was the audiotape recording of the teacher's responses to the post-task interview. The score was calculated as the number of individual children whose reactions to the story or to their personal experiences were described by the teacher.

Developing Children's Sense Perception

The work sample data for the Sense Perception module competencies included two audiotape recordings from each classroom visit and a third audiotape recording from visits on the second and fourth occasions. These were recordings of teachers' presentations of sense perception activities (either a set of beads or a set of spices) to a small group of children in the classroom, teachers' responses to an interview regarding the sense perception experiences, and teachers' generation of a list of possible uses of a set of colored pencils in activities for young children. The instruments used for rating the data produced a score for each of the three competencies of the module: Awareness, Diversity, and Diagnostics.

Awareness: The purpose of this measure was to determine the extent to which teachers were able to generate possible sense perception activities using a set of pencils. The pencils were selected for having potential use in both sense perception and non-sensory types of activities. They varied in shape, color, texture, and size. Teachers were told to generate as many activites as possible. They were not told

that the activities had to be sensory oriented. Each activity suggested by a teacher was rated as pertaining to one of the five senses or as providing a non-sensory experience. The Awareness score was calculated as the total number of sensory uses.

Diversity: The purpose of this measure was to determine how many of twenty categories of criterion teacher behavior were performed during the course of the recorded activity with the set of sensory materials. The twenty categories included five teacher behaviors which foster sense perception skills in each of four sense modalities. The five categories of teacher behavior included:

1. Make Statement. Teacher focuses attention by making a statement or comment (e.g., "This is a red one;" "Touch this.").
2. Ask Label. Identify, name, describe: Teacher asks child for verbal description or attribute label of single object (e.g., "How does it feel?" What color is it?").
3. Ask Relationship. Compare/contrast, discriminate, match: Teacher asks child to evaluate relationship between one object and others in the array (e.g., "How is this one different?" "Which one is red?" "Which is more rough?" "Put this one with the one that is the same.").
4. Ask Pattern. Group, classify, sort, order: Teacher asks child to arrange several objects in a pattern (e.g., "Put the rough ones here;" "Put the ones together that are the same; why do you think they are the same?" "Place these objects in a line in order from hard to soft.").

5. Ask Science. Explain, predict: Teacher asks child for a scientific rationale (e.g., "How would it look if we stepped on it?" Why is it red?").

The five sense modalities included Touch, Hearing, Sight, Smell and Taste. The Diversity score was computed as the number of categories with entries. Thus, the maximum possible score was twenty.

Diagnostics: The purpose of this measure was to indicate the extent to which a teacher produced a description of the level of skill or involvement of individual children in the sense-perception experience. The score was calculated as the number of individual children whose levels of skill and/or involvement were described by the teacher.

Rater Training

A total of six research assistants were withdrawn from data collection for data rating. A pair of research assistants was trained on the competencies of each module. Their training extended over a two week period and included five stages:

1. Reading selected pages of the module.
2. Taking the paper-pencil assessment instrument and discussing their responses.
3. Practicing a set of rating procedures using at least two sets of data collected during the dress rehearsal of the data collection training.
4. Developing and agreeing on a standard set of definitions.

5. Independently rating at least three new sets of dress rehearsal work sample data and then independently re-rating the same data. They achieved better than an 85% level of both inter- and intra-rater agreement on each measure.

Data Assignment

In order to estimate the extent to which inter-rater and intra-rater agreement was maintained during rating of actual data, the data from a subsample of eight trainees were selected for multiple ratings on each work-sample instrument. This subsample consisted of one high education and one low education trainee from each of the four treatment groups. Their data was rated twice by each rater. In addition, each rater was randomly assigned the data of half of the 23 subjects not in the subsample. The data to be rated was given to each rater as a randomly ordered list.

APPENDIX B

RELIABILITY FOR WORK SAMPLE AND PAPER-PENCIL MEASURES

GENERALIZABILITY

WORK SAMPLE MEASURES

GENERALIZABILITY

Generalizability analyses were performed on the ratings of all measures obtained from the sub-sample of eight teachers in order to determine the extent to which observed scores could be generalized across those factors which varied for each work sample method. Table B1 reports the mean squares for each facet of methodological variance. The components of variance are combined and expressed as a proportion of the observed score variance.

The generalizability analyses for the three physical arrangement measures (balance, quality, and integration) were based on fully crossed designs involving eight trainees, two occasions, and two raters.

The generalizability analyses of the measure Sense Perception awareness was based on a fully crossed design involving eight trainees and two raters.

The analyses of Literature Experience Measure, selection, was based on a crossed design of eight trainees and two occasions. The measure was scored objectively and had no raters.

The generalizability analyses of the measures, LE diversity and SP diversity and diagnostics, were based on partially nested designs. Raters, occasions, and persons were completely crossed. Materials were nested within occasions.

The results of the generalizability analysis of the Literature Experience diversity measure indicated that scores were stable across raters and occasions, but varied considerably between the two books used as stimulus materials. Additional information obtained from anecodotal observations suggested that one book was more difficult than the other.

Table B1. Facets of methodological variance from the generalizability study of work sample measures.

		Facets of Methodological Variance					
Competency Module	Competency Measure	Raters M.S.	Occasions M.S.	Materials (as if crossed) M.S.	Residual M.S.	Proportion Error of Observed Score	Design X=Crossed :=Nested
Physical Arrangement	Balance	0.78	0.03	--	0.22	0.12	pxoxr
	Quality	3.13	1.06	--	14.76	0.16	pxoxr
	Integration	0.31	1.53	--	0.32	0.16	pxoxr
Literature Experience	Diversity	0.28	0.03	7.03	3.21	0.90	pxrx(m:o)
	Diversity-BP	2.25	--	--	0.68	0.08	pxr
	Diversity-SD	0.56	--	--	1.56	0.44	pxr
	Personalization	--	0.56	1.56	4.25	1.00	pxrx(m:o)
	Personalization 01-04	--	0.62	.03	0.85	0.47	px(m:o)
	Selection	--	0.57	--	4.28	0.41	pxo
Sense Perception	Diversity	4.50	28.125	4.50	13.26	0.94	pxrx(m:o)
	Diversity 01	3.06	--	--	5.44	0.37	pxr
	Diversity 02	1.57	--	--	5.93	0.32	pxr
	Diagnostic	0.00	0.13	1.13	2.09	0.79	pxrx(m:o)
	Diagnostic 01-04	--	1.10	.03	7.67	0.43	px(m:o)
	Awareness	0.56	--	--	0.56	0.02	pxo

On the basis of the generalizability data and the additional information, the design of the evaluation study was modified so as to analyze the data from each book separately.

The results of the generalizability analysis of the Sense Perception diversity measure showed very small variation between the two sets of sensory materials or between the two raters. However, there was high variation between occasions one and two. Consequently, the design of the OFT evaluation study was modified so as to analyze the data from each occasion separately.

The results of the analysis of the Sense Perception diagnostic measure showed no variance due to the two raters. Based on the results of the analysis of rater variance and the similarity of the teacher behaviors measured, the number of double ratings was reduced by half for LE personalization.

Neither of these two measures were stable across occasions or materials at the pretraining administration. After the posttraining administration a generalizability analysis was done for both SP diagnostic and LE personlization across all four occasions using the data of the eight Control Group subjects. The results indicated that both measures showed some stability across occasions. The final analysis, however was done using each occasion.

Table B2 reports the overall generalizability coefficients for the measures, as they were used in the final analysis.

Table B2. Generalizability coefficients for work sample measures.

Competency Module	Competency Measure	Possible Score	Generalizability Coefficients (n=8)
Physical Arrangement	Balance	9	0.88
	Quality	80	0.84
	Integration	18	0.84
Literature Experience	Selection	10	0.59
	Diversity-BP	12	0.92
	Diversity-SD	12	0.55
	Personalization 01-04	10	0.53
Sense Perception	Awareness	(Frequential)	0.98
	Diversity 01	20	0.73
	Diversity 02	20	0.78
	Diagnostic 01-04	10	0.58

RELIABILITY

PAPER-PENCIL MEASURES

Reliability

Reliability analysis was done on paper-pencil measures to determine stability across occasions. Table B3 shows the test retest rank order correlation coefficients for the Control Group. It is clear from this analysis that test revision is warranted for selected subscales. An item analysis modification with plans for pilot testing are currently in progress and are expected to be completed for use in the coming year.

Two of the paper-pencil tests were made up in part of open-ended items Raters were trained to an 85 percent agreement criterion over 20 trials before rating the OFT data. All open-ended items were scored by two raters. Where they did not agree the average of their two scores was used for the final analysis. Table B4 shows the combined pre- and posttraining percent rater agreement for the following paper-pencil measures: PA total score, quality, integration, SP total score and all SP subscales. There were no open-ended items on the LE paper-pencil test.

Table B3. Reliability coefficients for paper-pencil measures.

Competency Module	Competency Measure	Possible Score	Test-Retest Coefficient (n=8)
Physical Arrangement	Balance	15	-0.23
	Quality	14	0.54
	Integration	11	-0.25
	Total Test	31	0.21
Literature Experience	Diversity	14	0.02
	Personalization	3	0.21
	Selection	6	-0.30
	Total Test	21	0.16
Sense Perception	Diversity	14	0.43
	Diagnostic	14	0.38
	Awareness	14	0.90
	Total Test	33	0.67

Table B4 Percentage rater agreement for open-ended paper-pencil test items.

Competency Module	Competency Measure	Possible Points (Open Items Only)	Percentage Rater Agreement (n=8)
Physical Arrangement	Balance	0	NA
	Quality	7	86.3
	Integration	3	83.9
	Total Test	7	86.3
Sense Perception	Diversity	5	88.8
	Diagnostic	3	83.9
	Awareness	7	83.9
	Total Test	15	85.5

APPENDIX C

SUPPLEMENTARY DATA: METHODS AND RESULTS

Supplementary Data

This appendix reports data obtained by four methods other than either the paper-pencil assessment instruments or the ratings of teachers' work samples. The instruments used to collect this data were developed in the process of conducting the evaluation study. They were not measures which had been specified in the original research design. They were developed in response to questions which arose during data collection. Their data are included here as supplementary information for interpreting the results of the research study

Trainee Feedback Questionnaire

After the completion of training, feedback questionnaires were mailed to each participating trainee. They requested ratings of (a) trainer effectiveness, (b) module effect (SP) and (c) characteristics of the children each trainee worked with. The forms had 29 common items and either 6 or 7 items specific to the unit. The responses indicated variation in several aspects of trainer performance and in trainees' working situation. Anonymous rating forms were returned by thirteen of the 23 trainees who participated in the research study. The responses came from four LE trainees, three SP trainees, and six PA trainees. A copy of this instrument can be found at the back of this appendix.

Working Situations

Different trainees were working in situations that varied considerably in class size and child characteristics. Class size varied from small as 11 to as large as 35. Children's age varied from 2-5 years with some

trainees having mixed ages while others had only 2 and 3 years olds and others only 4 and 5 years old. The number of bilingual children in a class varied from none to only one. The sex balance varied from equal numbers to twice as many of one than the other sex.

Trainer Performance

Variations in trainer performance may have served as intervening variables which influenced the outcomes of training.

The items assessing trainer effectiveness 1-13 were scored (0-4). The minimum, maximum, and average total scores for each trainer are reported in Table C1.

The scores indicate that the PA trainer was rated considerably higher than the other two.

Table C1. Total scores on trainee feedback questionnaire.

TRAINERS	MINIMUM	MAXIMUM	AVERAGE
PA (n=6)	31	50	40
SP (n=3)	19	31	27
LE (n=4)	17	40	29

The average scores for the 18 items assessing trainer and module effectiveness are presented in Table C2. There were six quantitative aspects of trainer performance and one quantitative rating of module performance.

Aspects of Trainer Performance:

1. Physical arrangement of the training environment.

   The PA workshop was the only one for which respondents rated the comfort of the seating as "excellent."

2. Knowledge of training materials.

   The trainees' responses to the open-ended questions and their ratings on item 9 suggest that the LE and PA trainers had better knowledge of the materials than the SP trainer.

3. Organization of content of training.

   Both LE and SP trainees wrote about problems with the pace, or tempo, of training. E.G., the LE course "moved much too slowly;" it was "dull and repetitive." The SP trainer "often finished earlier than she had planned and didn't know what to do so read a little more out of the book before letting the class go." This was also evident from the ratings of number of class sessions (Item 2) length of class meetings, (Item 3 as well as on the ratings of whether the trainer was well organized (item 11). Two of the three SP respondents indicated that the meetings were too long. One of these wrote in: "The teacher didn't have enough prepared most of the time."

Table C2. Average item scores on trainee feedback questionnaire.

DID YOUR TRAINER:	LE (N=4)	SP (N=3)	PA (N=6)
1. Provide materials you needed during class?	3.5	3.67	3.8
2. Explain difficult concepts and terms in your competency unit manual?	2.5	2.67	3.67
3. Allow opportunities during class time for discussion of material and ideas?	3.75	2.67	3.67
4. Ask questions that required you to think hard about the material you studied in class?	2.25	1.3	4.0
5. Help you discover ways to apply what you were learning?	2.3	1.67	4.3
6. Help you discover new activities to supplement what you studied in class?	1.25	1.0	3.2
7. Help you discover additional readings to supplement what you studied in class?	.25	0.3	1.8
8. Present material in class in the same order that it was in your work manual?	2.5	4.0	2.2
9. Know well the materials presented in each session?	3.25	2.0	3.5
10. Accept and value other's opinions?	2.5	2.67	4.0
11. Seem well-organized during class?	2.0	2.0	4.0
12. Seem enthusiastic and motivated during class?	1.75	1.3	3.3
13. Cover the right amount of material for you in each class session (Not too much, not too little)?	1.75	1.67	2.8
DURING YOUR TRAINING:			
1. Were there comfortable seating arrangements for you?	1.75	1.0	4.0
2. Was there the right number of class sessions? Too few, too many.	2.0	1.0	2.4
3. Was the length of the class meeting appropriate (not too long, not too short)?	2.25	0.7	2.75
4. Rate the teaching effectiveness of your trainer.	1.25	1.0	2.3
5. Rate the teaching effectiveness of your training manual.	1.25	2.0	3.3

4. Extending FLS training.

Four ratings items (5, 6, 7, & 8) suggest that the three trainers varied in their skill at extending beyond the content of the class and the module. (Note that for Item 8, 3 may be the optimal rating because a rating of 5 may indicate an inability to modify the information to suit the trainees. For example, one SP respondent gave the SP trainer a rating of 5, as Always presenting material in class in the order that it was presented in the module, and then wrote in: "Simply followed manual").

5. Motivation.

Three rating items differentiated between the PA trainer and the other two. The PA trainer was rated higher than the other two on Item 4 which might be characterized as "demand for excellence". Item 10 is a rating of trainer acceptance and valuing of others' opinions and item 12 rates enthusiasm and motivation during class. Again, the PA trainer outranked the LE and SP trainers.

6. Trainer and Module effectiveness.

The ratings for items 4 effectiveness of trainer and 5, effectiveness of module both produced the samerank ordering: PA, LE, SP. However, the module was rated at a higher level than the trainer for both the PA and SP units.

Group Interview

At the completion of data collection, a half-day meeting was held in San Diego to discuss the OFT participants reaction to the

training workshops and research study. Participants at the meeting included two members of the Laboratory staff, the trainers of the PA and LE workshops, and five teachers who had participated in training. The three training workshops were discussed in order PA, LE, SP.

Physical Arrangement Training

A participant in the PA workshop reported that all but the first class meeting has been held in her home. It was reported that at the beginning of training, many of the trainees had difficulty in implementing changes in their classrooms. They would arrive at school to find that the changes they had made the previous day had been undone by the alternate shift of teachers. However, by the sixth week, many physical arrangement changes had been established because the alternate teachers (2 to 3 teachers per class) had found that "this works." The children got along better, and the teachers did not need to work as hard. The physical environment was doing the work for them.

Several participants suggested a remedy for resistance to classroom change: All teachers of a classroom should take the workshop together. They agreed that the best time of day for such a class would be during the children's naptime--from lunch till 2:00 p.m.

Regarding the module, a participant noted that at first, the book had appeared to be aimed only at inexperienced teachers. However, she reported that considerable enthusiasm had developed as the trainees worked together with the module and one another.

The PA workshops apparently had followed a group problem-solving format, so that each participant became very familiar with the classrooms of other participants and cooperatively identified

improvements. It was suggested that the major area of improvement for participants was in their balance of different types of areas, however, this was frequently constrained by small room size. One participant noted that she was already doing what the module prescribed in terms of storage and accessibility of materials. However, she reported that accessibility of materials to child use was a controversial issue in the group.

This participant also shared an anecdote regarding her own application of module skills. Since her participation in the PA workshop, she had become involved in the planning of a new day care center. She was demonstrating her competence at a prespecified objective of the unit by incorporating into her plan the basic idea of the division of a center into learning areas arranged to accomplish particular objectives. But she had taken a <u>further</u> step in applying her learning. When she found that a new California law prohibited the use of "pits" and "lofts" because they are not possible for handicapped children to maneuver, she went beyond the unit's specific suggestion to an application of the underlying principle. She invented a new alternative. She was planning to build a movable platform on wheels which might serve the same purposes, but have the additional advantage of flexibility. She was extending and elaborating on the ideas she would gained from the workshop. She stated emphatically that the PA module should be taught first in a sequence of FLS teacher training workshops.

<u>Literature Experiences Training</u>

One participant in the LE workshop who had also received trainer training reported that in her experience, teachers already plan

lessons around stories, so that the module content was not new.

Another participant reported that her expectations for the module had been disappointed, because she had hoped for "new knowledge," that is, a how-to-do-it type of information. She said that the LE trainees were "just starting to get comfortable as a group, when it was time to leave, I was ready to start, and it ended." She reported that for her, the most important part of the training was the group's development of book selection guidelines. She appreciated being able to consider specific features in selecting books, rather than her prior general impressions.

A third LE trainee reported that the group discussions were more important than reading the manual.

Members of the LE group indicated that the LE module served as a "refresher" for the trainees with B.A. degrees and this enabled them to improve more than the trainees with high school diplomas. That is, the module may have reminded teachers to practice what they had learned elsewhere.

Sense Perception Training

One SP trainee said that "once you had gone through it (generating word lists, etc.) a couple of times, additional activities were unnecessary." The unit was too long, too wordy, too redundant. There was nothing new added to our suggested interpretation of the SP data curves, i.e., high education trainees were bored by the redundancy; low education trainees were enabled to learn better with the redundancy.

Trainer Interview

The Physical Arrangement and Literature Experience trainers were interviewed following the OFT study. They were questioned about their experiences as Trainers and their opinions regarding the OFT training process and research study. The Sense Perception trainer was not available for the interview. The trainer of the Physical Arrangement workshops was the director of the Head Start Program and had previous training experiences---including limited use of FLS competency modules. He made some general comments about the usefulness of the materials and the appropriateness of the evaluation study measures. He repeated several times something that was obviously important to him: "These materials can enable implementation of an open approach." However, he also saw a danger in use of the modules: They could be used prescriptively, dogmatically,---"This is the way you do it." If the modules are taken as a final right, or expert, answer, they are being misused. Each module provides no more than an initial framework for thinking about and organizing a particular aspect of early education. The important outcome is how well teachers internalize and make personal application in their own situation.

The FLS staff agrees wholeheartedly with the PA trainers perceptions of the FLS. The modules are only a tool and can be used for a variety of purposes. There are two stages of application of training. First is to take a set of rules and apply them to a situation. Second, is to take the underlying principles, and apply them in the generation of a personal set of working hypotheses.

The learning model underlying the design of our competency modules is that competence at a given set of module-prescribed rules is a necessary step toward competence at developing a set of personal rules. The important point is that the FLS has the potential of achieving the second set of personal rules and the danger of being limited to the first set of prescribed rules.

In regard to measurement for evaluation, the PA trainer noted that the responsive model is open-ended." He argued that the measurement should be correspondingly open-ended; that is, outcomes should be assessed not only in terms of pre-specified objectives, but also in terms of the ways and extent to which people make diverse application of their learnings.

In our conversation with the trainer of the LE workshops, it appeared that this was her first training experience (this may account for the slow development of group dynamics in her workshop). She reported that some sections of the module were difficult for her. She also said that the organization of the manual was not clear, and had felt compelled to the order during her training workshop. As one example, she mentioned that it was confusing to have the book selection guidelines adjacent to the criteria for selecting books to enhance self-concept. As another example, she reported that she had presented the conceptual table at the end of training because it had been confusing to her when presented at the beginning during trainer-training. She said that she considered the most important competency taught in the module to be "organizational principles." From her explanation, it appeared that she meant something like skills at organizing complex information.

Trainer Ratings of Trainees

At the conclusion of the trainer interview, the trainers of the Literature Experiences and Physical Arrangement modules were asked to rank order their trainees according to level of competence. Table C3 shows the relationships between their orderings and their trainees' performance on the paper-pencil and work sample measures.

The table suggests that the trainer ratings of the Physical Arrangement trainees' competence was positively related to their work-sample ratings. This ability to diagnose trainee's level of competency at arranging the Physical Environment suggests that the PA trainer possesses an ability to personalize his training, that is, to be aware of individual strengths and weaknesses of his trainees. Note that the Literature Experience trainer was not overly successful at matching trainees with their actual performance. This trainer was also rated as less than satisfactorily effective by LE trainees. Thus, the ability to personalize may be a critical one for FLS trainers to possess or develop.

Research Participation

The meeting ended with a discussion of the trainees' reaction to the conduct of the research study. Most of the trainees present took the pre- and posttraining paper-pencil tests matter-of-factly. However, couple of trainees suggested that it was a negative way to start the first day of training.

In regard to the procedures of performance assessment, the trainees commented favorably on the visits of the research assistant

Table C3. Rank order correlations between trainers' ratings of trainees' level of competence and trainees' performance on paper-pencil and work sample measures at the conclusion of training.

Physical Arrangement Competencies		Rank Order Correlations with PA Trainers' Ratings
Paper-Pencil	Balance	0.34
	Quality	0.52
	Integration	-0.01
	Total Score	0.68
Work Sample	Balance	0.47
	Quality	0.73
	Integration	-0.48
Literature Experience Competencies		**Rank Order Correlations with LE Trainers' Ratings**
Paper-Pencil	Diversity	-0.21
	Personalization	0.06
	Selection	0.29
	Total Score	0.26
Work Sample	Diversity-BP	0.48
	Diversity-SD	0.14
	Personalization-01	-0.06
	Personalization-02	0.00
	Selection	-0.62

as persons; they liked them. However, they objected to two things, a) lack of prior orientation to the general purpose and nature of participant involvement in the research study, and b) insufficient information regarding the purpose of the specific classroom performance tasks. One participant recommended that if similar assessment were to be included in a training program, the information should be used to provide feedback to trainees.

In summary, the participants at the meeting were open and frank in their comments and generally positive, if not enthusiastic, about their participation in training.

Anecdotal Observations

After the completion of data collection, all research assistants met with the evaluation study staff. They used the note sheets from each visit to recall and share anecdotes and opinions related to the effects of training and the effects of research study participation. The tape recordings of this meeting were reviewed, and a transcript was made of comments specific to each trainee. The data collectors' anecdotes revealed differences in site characteristics, suggested events which may have been confounded with the effects of treatment, and provided examples of the effects of training on the module: Physical Arrangement of the Classroom. As the data collectors had been instructed to leave the area during the recording of the literature and sense perception experiences, they reported no anecdotes related to the effects of training with the other two modules.

Site Differences

The research assistants' descriptions of the settings in which the trainees were working with children indicated considerable variation in size and number of rooms. Some trainees worked in small single rooms, some worked in large settings with responsibility for several rooms, and some worked in large settings with responsibility for only one room. For the conduct of the sense perception and literature experience activities, some teachers took their small groups of children into separate, private rooms while others conducted their experiences among the ongoing classroom activities. The number of other adults also available to work with the children varied. Some trainees had several assistants and volunteers, and a few even had substitute teachers assume classroom responsibilities during data collection. Other trainees assumed sole responsibility for working with their groups of children while concurrently managing the remaining children in the classroom.

Intervening Events

During the six week training period events other than treatment were reported which possibly affected trainee performance. For example, two trainees in the research study Control group participated in a workshop on classroom arrangement. Also, on the last visit, two trainees in the LE group were sharing their space with a class of 30 additional children and their teachers whose own classroom had been torn down. The research assistant who visited these teachers reported them both to be calm and cooperative in spite of the difficulties of the additional children.

Training Effects

The anecdotes reported about the trainees in the Physical Arrangement

group provide some insight into the performance and assessment of teaching competencies related to arranging a classroom. The two trainees who were rated as least competent by the PA trainer were reported as having identical room arrangements -- apparently under the direction of their common administrator. Their comments led the visitor to conclude that they were quite satisfied with their arrangements on the basis of being able to see every child from the center of the class.

Positive changes in the physical arrangement were reported for four of the PA trainees. Posttraining observer comments included the following:

> Teacher #21. Occasion 3.
>
> .....Definite change in physical arrangement. Each area of the classroom was much more clearly defined in the teacher's mind and also in the way it was arranged. She was very concerned that the materials were available for the kids.
>
> Teacher #29. Occasion 3.
>
> .....She couldn't stop talking about the class, loved it, said she would change her room around several times. All the other teachers were upset with her for doing that at first, and then they got used to it and started enjoying the room arrangement.
>
> Teacher #07. Occasion 4.
>
> .....She did <u>everything</u>. Between the third and the fourth... was about 5 days...she changed everything around. She said she finally understands it -- what we are asking her. She did not understand it before.

Biographical Data Card

FAR WEST LABORATORY
FOR EDUCATIONAL RESEARCH AND DEVELOPMENT

FLEXIBLE LEARNING SYSTEM
Biographical Card - All Participants

NAME _____
(Last, First, Middle Initial)

HOME ADDRESS _____
(Street, City, State, Zip Code)

TELEPHONE _____
(Area Code, Number)

LEARNER I.D. NUMBER
TRAINER I.D. NUMBER
SITE I.D. NUMBER
WORK UNIT I.D. NUMBER
LEARNING UNIT I.D. NUMBER

TRAINER _____

JOB TITLE
1 [ ] *teaching assistant (aide)*
2 [ ] *teacher*
3 [ ] *coordinator/trainer*
4 [ ] *administer*
5 [ ] *other (Please specify)*

[6] [ ] *not teaching presently*

CHILDREN YOU WORK WITH
1 [ ] *pre-kindergarten*
2 [ ] *K-3rd grade*
3 [ ] *other (Please specify)*

FORMAL EDUCATION *Yrs Compltd.*
1 [ ] *6 yrs or less (elem.)*
2 [ ] *12 yrs or less (high school)*
3 [ ] *14 yrs or less (jr. college)*
4 [ ] *16 yrs or less (college)*
5 [ ] *17 yrs or more (grad. prof. school)*

HIGHEST DEGREE ATTAINED
1 [ ] *high school diploma*
2 [ ] *A.A.*
3 [ ] *B.A.*
4 [ ] *M.A.*
5 [ ] *Ph.D.*

DO YOU HAVE A TEACHING CREDENTIAL?
1 [ ] *yes*
2 [ ] *no*

YRS OF WORKING EXPERIENCE IN EDUCATION
1 [ ] *one year or less*
2 [ ] *3 years or less*
3 [ ] *8 years or less*
4 [ ] *15 years or less*
5 [ ] *over 15 years*

AGE
1 [ ] *under 30*
2 [ ] *30 - 45*
3 [ ] *over 45*

SEX
1 [ ] *male*
2 [ ] *female*

ARE YOU A PARENT?
1 [ ] *yes*
2 [ ] *no*

HOW DO YOU PLAN TO USE THIS UNIT?
1 [ ] *teach in classroom*
2 [ ] *train other adults*
3 [ ] *other (Please specify)* _____

Note to the Learner -

Here is a set of items that will give you an opportunity to rate several aspects of the training course you recently completed:

1. Your feelings about the course.
2. The trainer you worked with.
3. The materials you worked with.

In addition, there is a set of questions about the children you worked with.

Your trainer knows that you are filling out this instrument. He/she will not see your ratings. It will take you approximately 15 - 20 minutes to answer the items, and your answers will help us determine what is happening during training. Please answer as honestly and objectively as possible and return this form in the enclosed envelope to Terry D. Taylor at Far West Laboratory.

1. Describe your feelings and expectations about taking this course before it began.

2. Describe your feelings and expectations about the course after you finished it.

3. What parts of the training were most meaningful to you?

4. What were the weaknesses of the training?

Check one of these boxes for each question.

DURING YOUR TRAINING:

	Unsatis-factory			Satis-factory	Excel-lent
	1	2	3	4	5
1. Were there comfortable seating arrangements for you?					
2. Was there the right number of class sessions? Too few, too many.					
3. Was the length of the class meetings appropriate (not too long, not too short)?					

If you checked 1 or 2, for question #3, was it:

[ ] Too long. [ ] Too short.

	Unsatis-factory			Satis-factory	Excel-lent
	1	2	3	4	5
4. Rate the teaching effectiveness of your trainer.					
5. Rate the teaching effectiveness of your training manual.					

DID YOUR TRAINER:

	Never		Some-times		Always
	1	2	3	4	5
1. Provide materials you needed during class?					
2. Explain difficult concepts and terms in your competency unit manual?					
3. Allow opportunities during class time for discussion of material and ideas?					
4. Ask questions that required you to think hard about the material you studied in class?					
5. Help you discover ways to apply what you were learning?					
6. Help you discover new activities to supplement what you studied in class?					
7. Help you discover additional readings to supplement what you studied in class?					
8. Present material in class in the same order that it was in your work manual?					
9. Know well the materials presented in each session?					
10. Accept and value other's opinions?					
11. Seem well-organized during class?					
12. Seem enthusiastic and motivated during class?					
13. Cover the right amount of material for you in each class session (Not too little, not too much)?					

If you checked 1 or 2, for question #13, was it: [ ] Too little. [ ] Too much.

DID YOUR TRAINER:

	Never		Some-times		Always
	1	2	3	4	5
1. Provide an array of books for you to work with in class?					
2. Encourage you to select books to use with children from this array?					
3. Encourage you to work in groups during class sessions?					
4. Ask you to select your favorite books from the array?					
5. Encourage you to bring your own favorite children's books to class?					
6. Provide clear examples of planning and facilitating responsive literature experiences?					

DID YOUR TRAINER:

1. Ask you to observe the physical arrangement of the classroom you work in?
   Check one: ☐ Yes ☐ No.

2. Ask you to describe your classroom to the rest of the people in your training class?
   ☐ Yes ☐ No.

3. Ask you to prepare a floorplan of your classroom? ☐ Yes ☐ No.

4. If you did prepare a floorplan, did your trainer ask you to:
   a. measure the room and draw it to scale? ☐ Yes ☐ No.
   b. draw in the furniture in the room? ☐ Yes ☐ No.
   c. draw in the boundaries in the room? ☐ Yes ☐ No.

5. Have class discussions of the arrangement of classrooms? ☐ Yes ☐ No.

6. If there were discussions of classroom arrangement, were they accompanied by:
   a. floorplans? ☐ Yes ☐ No.
   b. classroom diagrams? ☐ Yes ☐ No.
   c. photographs? ☐ Yes ☐ No.
   d. other (please specify)? ☐ Yes ☐ No.

7. What did you discuss about the classrooms?
   a. How the arrangement and location of different areas realte to one another? ☐ Yes ☐ No.
   b. How the inside of an area is arranged? ☐ Yes ☐ No.
   c. How to relocate or rearrange an area? ☐ Yes ☐ No.
   d. other (please specify):_____

_____

_____

DID YOUR TRAINER:

	Never		Some-times		Always
	1	2	3	4	5
1. Provide experiences that helped you understand how your senses function?					
2. Explain how children learn spontaneously through their senses?					
3. Help you learn how to use your manual to define cognitive skills as underlying objectives for sensory experiences?					
4. Help you learn how to use your manual to make sensory experience objectives easier or harder?					
5. Help you learn how to make sensory games out of everyday materials?					
6. Help you learn how to use the diagnostic skills presented in your manual for evaluating a child's level of involvement?					

The following set of questions will provide a description of the children in your class. This will help us know what kinds of children we should be considering when we develop competency manuals.

1. What is the age range of children in your class (for example, 3-5 year-old children)?

2. Are there any children that are physically handicapped? How many?

3. How many boys in your class?

4. How many girls?

5. Are there any bilingual children in your class? How many?

6. How many new children have entered your class since January?

7. Describe what effects you think your training might have had on the children you worked with.